

This textbook investigates definiteness both from a comparative and a theoretical point of view, showing how languages express definiteness and what definiteness is. It surveys a large number of languages to discover the range of variation in relation to definiteness and related grammatical phenomena: demonstratives, possessives, personal pronouns. It outlines work done on the nature of definiteness in semantics, pragmatics and syntax, and develops an account on which definiteness is a grammatical category represented in syntax as a functional head (the widely discussed D). Consideration is also given to the origins and evolution of definite articles in the light of the comparative and theoretical findings. Among the claims advanced are that definiteness does not occur in all languages though the pragmatic concept which it grammaticalizes probably does, that many languages have definiteness in their pronoun system but not elsewhere, that definiteness is not inherent in possessives, and that definiteness is to be assimilated to the grammatical category of person.

CAMBRIDGE TEXTBOOKS IN LINGUISTICS

General editors: s. r. anderson, j. bresnan, b. comrie, w. dressler,
c. ewen, r. huddleston, r. lass, d. lightfoot, j. lyons,
p. h. matthews, r. posner, s. romaine, n. v. smith, n. vincent

DEFINITENESS

In this series

- p. h. matthews *Morphology* Second edition
- b. comrie *Aspect*
- r. m. kempson *Semantic Theory*
- t. bynon *Historical Linguistics*
- j. allwood, l.-g. anderson and ö. dahl *Logic in Linguistics*
- d. b. fry *The Physics of Speech*
- r. a. hudson *Sociolinguistics* Second edition
- a. j. elliot *Child Language*
- p. h. matthews *Syntax*
- a. radford *Transformational Syntax*
- l. bauer *English Word-Formation*
- s. c. levinson *Pragmatics*
- g. brown and g. yule *Discourse Analysis*
- r. huddleston *Introduction to the Grammar of English*
- r. lass *Phonology*
- b. comrie *Tense*
- w. klein *Second Language Acquisition*
- a. j. woods, p. fletcher and a. hughes *Statistics in Language Studies*
- d. a. cruse *Lexical Semantics*
- f. r. palmer *Mood and Modality*
- a. radford *Transformational Grammar*
- m. garman *Psycholinguistics*
- w. croft *Typology and Universals*
- g. g. corbett *Gender*
- h. j. giegerich *English Phonology*
- r. cann *Formal Semantics*
- p. j. hopper and e. c. traugott *Grammaticalization*
- j. laver *Principles of Phonetics*
- f. r. palmer *Grammatical Roles and Relations*
- b. blake *Case*
- m. a. jones *Foundations of French Syntax*
- a. radford *Syntactic Theory and the Structure of English: a Minimalist Approach*
- r. d. van valin, jr. and r. j. lapolla *Syntax: Structure, Meaning and Function*
- a. duranti *Linguistic Anthropology*
- a. cruttenden *Intonation* Second edition
- j. k. chambers and p. trudgill *Dialectology* Second edition
- c. lyons *Definiteness*

DEFINITENESS

CHRISTOPHER LYONS

lecturer in linguistics
university of cambridge



CAMBRIDGE
UNIVERSITY PRESS

published by the press syndicate of the university of cambridge
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

cambridge university press

The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>
40th West 20th Street, New York, NY10011-4211, USA <http://www.cup.org>
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1999

This book is in copyright. Subject to statutory exception and to the provisions
of relevant collective licensing agreements, no reproduction of any part may
take place without the written permission of Cambridge University Press.

First published 1999

Printed in the United Kingdom at the University Press, Cambridge

Typeset in Times 9/13 [gc]

A catalogue reference for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Lyons, Christopher.

Definiteness / Christopher Lyons.

p. cm. — (Cambridge textbooks in linguistics)

Includes bibliographical references and index.

ISBN 0 521 36282 2 (hardback). — ISBN 0 521 36835 9 (paperback)

1. Definiteness (Linguistics) 2. Grammar, Comparative and
general – Syntax. 3. Semantics. 4. Pragmatics. I. Title.

II. Series.

P299.D43L97 1998

415—dc21 98–24724 CIP

ISBN 0 521 36282 2 hardback

ISBN 0 521 36835 9 paperback

*To the memory of my parents,
Edith and Patrick Lyons*

CONTENTS

<i>Preface</i>	<i>page</i> xv
<i>Abbreviations</i>	xviii
1 Basic observations	1
1.1 What is definiteness?	1
1.1.1 Simple definites and indefinites	1
1.1.2 Familiarity and identifiability	2
1.1.3 Uniqueness and inclusiveness	7
1.1.4 Identifiability, inclusiveness and indefinites	12
1.1.5 A unified account?	13
1.2 Types of definite noun phrase	15
1.2.1 Testing for definiteness	15
1.2.2 Demonstratives	17
1.2.3 Proper nouns	21
1.2.4 Possessives	22
1.2.5 Personal pronouns	26
1.2.6 Universal quantifiers	32
1.3 Simple and complex indefinites	33
1.3.1 Indefiniteness and cardinality	33
1.3.2 Complex indefinites	36
1.4 The noun phrase	41
1.5 Definiteness beyond the noun phrase	45
2 Simple definites and indefinites	47
2.1 Definiteness and indefiniteness marking	47
2.1.1 The occurrence of articles in languages	48
2.1.2 The distribution of definites and indefinites	51
2.2 Definites: semantic content and behaviour	52
2.2.1 Optional definite articles	52
2.2.2 Anaphoric articles	53
	ix

Contents

2.2.3 Deictic distinctions	55
2.2.4 Definite and specific	57
2.2.5 Nominalizing and other functions	60
2.3 Definite article types	62
2.3.1 Free-form articles	63
2.3.2 Bound articles	68
2.3.3 Phrasal clitics	72
2.3.4 Mixed systems and double determination	77
2.3.5 Definite adjectives	82
2.4 Other ways of expressing definiteness	86
2.4.1 Adpositional marking	86
2.4.2 Agreement	86
2.4.3 Pronominal marking	88
2.4.4 Word order	88
2.5 Indefinites	89
2.5.1 Indefinite article	89
2.5.2 Quasi-indefinite article	95
2.5.3 Types of cardinal article	100
2.5.4 Partitive indefinites	100
2.5.5 Bare indefinites	103
2.6 General remarks	105
 3 Complex definites and indefinites	 107
3.1 Demonstratives	107
3.1.1 Deictic distinctions	107
3.1.2 Non-deictic distinctions	113
3.1.3 Forms and positions	116
3.2 Proper nouns	121
3.3 Possessives	124
3.3.1 Pronoun and full noun phrase possessives	124
3.3.2 Affixal possessives	125
3.3.3 Alienable and inalienable	128
3.3.4 Possessives and definiteness	130
3.4 Personal pronouns	134
3.4.1 Semantic and grammatical content	134
3.4.2 Strong, weak and null pronouns	137
3.4.3 Personal determiners	141
3.4.4 Definite and demonstrative personal forms	145
3.5 Universals and other quantifiers	148

3.6	Indefinites	148
3.6.1	Full noun phrases	149
3.6.2	Pronouns	150
3.7	Indefinite demonstratives	151
3.8	Vocatives	152
3.9	Non-configurational languages	153
4	Some semantic and pragmatic distinctions	157
4.1	One definiteness or several?	157
4.1.1	Identifiability and inclusiveness	158
4.1.2	Anaphoric and non-anaphoric definites	158
4.1.3	“Deictic” and non-“deictic” definites	160
4.2	Specificity and referentiality	165
4.2.1	Opacity and scope ambiguities	166
4.2.2	Transparent contexts	170
4.2.3	A unified account of specificity?	172
4.3	Generics	179
4.3.1	Generics in English	181
4.3.2	Singular generics	184
4.3.3	Non-singular generics	189
4.3.4	Generics and proper nouns	193
4.4	Concluding comments	198
5	Interaction with other grammatical phenomena	199
5.1	Direct object marking	199
5.1.1	Differential case marking	200
5.1.2	Definite object marking	202
5.1.3	Oblique direct objects	205
5.2	Verb agreement	207
5.2.1	Direct object agreement	207
5.2.2	Indirect object agreement	211
5.2.3	Subject agreement	212
5.3	Definiteness and animacy	213
5.4	Articles and nominal feature marking	215
5.5	Null and implicit noun phrases	219
5.5.1	Pro	220
5.5.2	Traces	221
5.5.3	PRO	223
5.5.4	Implicit arguments	223

5.6 Classifiers	225
5.7 Concluding point	226
6 Definiteness effects	227
6.1 Discourse structure	227
6.1.1 The organization of information	227
6.1.2 Information structure and definiteness	232
6.2 Existential sentences	236
6.2.1 The phenomenon	237
6.2.2 Explanatory accounts	241
6.3 Other definiteness effects	246
6.3.1 Superlatives	246
6.3.2 PP-extraposition and <i>any</i> opacity	248
6.3.3 Property predication	249
6.3.4 Internal-head relatives	250
6.3.5 Concluding remarks	251
7 Defining definiteness	253
7.1 The grammatical, logical and pragmatic traditions	253
7.2 Hawkins: inclusiveness and location	260
7.2.1 Shared sets, inclusiveness and exclusiveness	260
7.2.2 Subsequent developments	262
7.3 Definiteness and quantification	265
7.3.1 Milsark and the quantification restriction	265
7.3.2 Generalized quantifiers	266
7.3.3 Discourse semantics: Heim	268
7.4 Relevance theory	271
7.5 Well, what is definiteness then?	274
7.5.1 Grammaticalization	275
7.5.2 Grammaticalization and definiteness	277
7.5.3 Simple and complex definiteness	279
8 Definiteness and noun phrase structure	282
8.1 The noun phrase as NP	282
8.1.1 The specifier position and definite Dets	283
8.1.2 Affixal articles	287
8.1.3 The definite constraint	288
8.2 The DP hypothesis	290
8.2.1 Movement processes in DP	291
8.2.2 Other functional categories in the noun phrase	294
8.2.3 Weaknesses in the DP analysis	296

8.3 A modified DP analysis	298
8.3.1 D as definiteness	298
8.3.2 The content of determiners	301
8.3.3 The analysis of personal pronouns	303
8.4 Configurationality and definiteness	305
8.5 Definiteness and person	310
8.5.1 Personal determiners and pronouns	310
8.5.2 Person and definite determiners	313
8.5.3 The nature of person-definiteness	318
8.6 Summary of proposals	320
 9 Diachronic aspects	 322
9.1 The emergence of functional structure	322
9.1.1 The development of DP	323
9.1.2 A definite article cycle	326
9.1.3 Indefinites and CardP	330
9.2 The origin of articles	331
9.2.1 Demonstrative to definite article	331
9.2.2 Articles in competition	334
9.2.3 Numeral to cardinal article	335
9.3 The longer perspective	336
9.3.1 Expansion in article use	336
9.3.2 The life cycle of definite articles	337
9.4 In conclusion	340
 <i>References</i>	 341
<i>Index</i>	359

PREFACE

This book is primarily a survey, but, unlike some other topic-based books in this series, it surveys two areas. First, it offers an account of the range of variation displayed by languages in relation to definiteness and related grammatical concepts. Most languages do not have “articles”, and in those that do they vary strikingly in both their form and their range of use. All languages have demonstratives, personal pronouns, possessives, and other expressions which either seem to be inherently definite or to interact in interesting ways with definiteness; but again, there is considerable variation in the ways in which these expressions relate to definiteness. Second, the book gives a (very selective) outline of the theoretical literature on definiteness. This literature is vast, consisting both of direct accounts of definiteness and of work mainly concerned with other phenomena on which definiteness impinges. Both the cross-linguistic survey and the theoretical survey are introductory and far from complete, and many of the choices I have made in reducing the material to manageable proportions are no doubt arbitrary. This is true particularly as regards the literature, where I have had to omit much which I see as important, and it is essential that the reader follow up further references given in the works I do refer to.

This is not just a survey, however. I am much too interested in the topic not to want to present my own view of what definiteness is, and I believe the work gains in coherence from the aim of reaching and defending (if in outline) a preferred account. Chapters 7 to 9, in particular, contain much discussion of the approach I believe to be the most promising. But in the earlier chapters too, I have not hesitated to advance far-reaching claims anticipating this approach. The view of definiteness I propose may be wrong, of course, but it will have achieved its purpose if a student reading my proposals is spurred to investigate further and show their inadequacy. My aim in this book is not to present a set of facts and analyses to be learned, but to offer a body of ideas to be thought about and improved upon.

The investigation of definiteness necessarily takes the reader into several domains of inquiry, some of which (like semantics and syntax) are highly technical. While I assume some familiarity on the part of the reader with the principles and

Preface

methods of linguistics, I do not assume advanced competence in these domains, and I have given at appropriate points brief outlines of essentials and references to further reading, where possible at an elementary level. But it must be stressed that the interested reader would need to follow up these references, sometimes to a fairly advanced point, in order to come fully to grips with the issues in question. I have in general maintained neutrality between different theoretical frameworks, except as regards syntax, where I assume the principles-and-parameters approach which is the most highly developed and best known. Most of the text of this book was written at a time when the current “minimalist” version of this approach was in its infancy and there were few accessible accounts available of this framework to refer the reader to, so I have taken little account of minimalism. But there is little in the syntax discussed here which cannot be easily recast in this paradigm.

There has been much debate over the years on the relative merits of, on the one hand, the wide-ranging descriptive work of typological studies, and, on the other, the deep analysis of a smaller range of languages done in theoretical work. I firmly believe that descriptive breadth and analytical depth benefit one another, but that the latter must be the ultimate goal, and I hope that the gulf between these two approaches to language is narrower now than it was. The “new comparative linguistics” in generative work indicates a recognition among theorists of the value of cross-linguistic investigation, though some of it can be criticized as too selective in scope. But, to repeat a familiar point, our understanding of the way language works is deepened by bringing to bear serious analyses of languages, not mere observational facts. And even the best descriptive grammars are rarely adequate by themselves to provide the basis for an analysis of any depth of a specific aspect of linguistic structure. Indeed, even the descriptive observations and generalizations made in typological work must be treated with great caution, partly because the descriptive grammars on which they are based are often unclear on crucial points or analytically unsophisticated, partly because the typologist looking at unfamiliar languages in pursuit of a generalization is prone to the same inaccuracy as the theorist aiming to prove a point of theory. In my own cross-linguistic survey here, I too will certainly have included inaccuracies, and I urge the reader to treat it as a guide and starting point, not as fully reliable data.

Many people have helped me in various ways in the course of my writing this book, and I wish to thank in particular the following friends and colleagues who have read and commented on the manuscript or sections of it, or discussed particular points with me: Nigel Vincent, Deirdre Wilson, Noel Burton-Roberts, Kasia Jaszczolt. Most special thanks to Ricarda Schmidt for constant intellectual and moral support.

A couple of points concerning the presentation of the material should be noted. Where items of literature discussed exist in different versions, I have tried to refer to the most easily accessible version. In the case of doctoral theses subsequently published this means the formal publication. The effect is sometimes that my reference is to a version dated several years later than the version most commonly cited. Finally, a note on my use of gender-marked personal pronouns in describing conversational exchanges: I follow the convention that the speaker is, unless otherwise stated, female, and the hearer or addressee male.

ABBREVIATIONS

1	first person
1EXC	first person exclusive
1INC	first person inclusive
2	second person
3	third person
ABL	ablative case
ABS	absolutive or absolute case
ACC	accusative case
ADESS	adessive case
ANA, ana	anaphoric
ART	article
ASP	aspect
ASS1, Ass1	associated with first person
ASS1EXC	associated with first person exclusive
ASS1INC	associated with first person inclusive
ASS2, Ass2	associated with second person
ASS3	associated with third person
AUX	auxiliary
C	common gender
CL	clitic
CLASS	classifier, class marker
CONT	continuous aspect
DAT	dative case
DECL	declarative
DEF, Def	definite
DEM, Dem	demonstrative
DIR	direction
DIST	distal
DU	dual number
ELAT	elative case

ERG	ergative case
EXP	experiential aspect
F	feminine gender
FUT	future tense
GEN	genitive case
GENR	generic aspect
HAB	habitual aspect
HON	honorific
IMP	imperative
IMPF	imperfect, imperfective aspect
IMPRS	impersonal
INAN	inanimate
INDEF	indefinite
INESS	inessive case
INST	instrumental case
INTR	intransitive
IRR	irrealis
LINK	linker
LOC	locative case
M	masculine gender
N	neuter gender
NEG	negative
NOM	nominative case
NONPAST	non-past tense
NUN	numeration
OBJ	object
OBL	oblique case
PART	partitive case
PASS	passive voice
PAST	past tense
PERS	person
PL	plural number
POSS	possessive
PRES	present tense
PRF	perfective aspect
PRI	primary case
PROX, Prox	proximal
PRT	particle
REFL	reflexive
REL	relative marker

Abbreviations

SG, Sg	singular number
SUBJ	subject
TNS	tense
TOP	topic
WH	interrogative

The standard labels are used for syntactic categories (N, V, D and Det, Agr etc.). Any idiosyncratic or non-standard labels used are explained at the appropriate point in the text.

1

Basic observations

This chapter sets the scene by presenting some basic issues and ideas, which will be investigated in greater depth in the rest of the study. It begins by examining the concept of definiteness itself, to establish a preliminary account of what this concept amounts to. This is followed by consideration of the various types of noun phrase which are generally regarded as definite or indefinite – since definiteness and indefiniteness are not limited to noun phrases introduced by *the* or *a*. Finally, some basic ideas concerning the syntactic structure of noun phrases are presented in outline. English is taken as the starting point, with comparative observations on other languages where appropriate, because it is easier and less confusing to outline basic issues as they are instantiated in one language, where this can be done, than to hop from one language to another. For this purpose, English serves as well as any language, since it has readily identifiable lexical articles, which make definite and indefinite noun phrases on the whole easy to distinguish. It is important to bear in mind that the discussion in this chapter is preliminary, and aims at a tentative and provisional account of the points examined. Many of the proposals made here and solutions suggested to problems of analysis will be refined as the study progresses.

1.1 What is definiteness?

I begin in this section by attempting to establish in informal, pre-theoretical terms what the intuitions about meaning are that correspond to our terming a noun phrase “definite” or “indefinite”.

1.1.1 *Simple definites and indefinites*

In many languages a noun phrase may contain an element which seems to have as its sole or principal role to indicate the definiteness or indefiniteness of the noun phrase. This element may be a lexical item like the definite and indefinite articles of English (*the*, *a*), or an affix of some kind like the Arabic definite prefix *al-* and indefinite suffix *-n*. I shall refer to such elements by the traditional label **article**, without commitment at this stage to what their grammatical status actually

is. Of course not all noun phrases contain an article – probably in any language – though the definite–indefinite distinction is never thought of as applying only to those that do. This is clear from the fact that in English *this house* would usually be judged (at least by linguists and grammarians) to be definite and *several houses* indefinite; judgments would probably be more hesitant over *every house*. Noun phrases with *the* and *a* and their semantic equivalents (or near-equivalents) in other languages can be thought of as the basic instantiations of definite and indefinite noun phrases, in that the definiteness or indefiniteness stems from the presence of the article, which has as its essential semantic function to express this category.¹ I shall refer to such noun phrases as **simple definites** and **simple indefinites**, and I limit the discussion to them in this section to avoid any possibility of disagreement over the definite or indefinite status of example noun phrases.

So the question we are concerned with is: What is the difference in meaning between *the car* and *a car*, between *the greedy child* and *a greedy child*, between *the hibiscus I planted last summer* and *a hibiscus I planted last summer*? Many traditional grammars would give answers like the following: *The* indicates that the speaker or writer is referring to a definite or particular car etc., not just any. But apart from being rather vague, this answer is quite inaccurate. If I say *I bought a car this morning*, I am not referring to just any car; the car I bought is a particular one, and is distinguished in my mind from all others. Yet *a car* is indefinite. There is in fact no general agreement on what the correct answer is, but two major components of meaning have been much discussed, and I introduce these in 1.1.2 and 1.1.3 in relation to some illustrative English data.

1.1.2 Familiarity and identifiability

Continuing with the example just considered, compare the following two sentences:

- (1) I bought **a car** this morning.
- (2) I bought **the car** this morning.

The car here is in some sense more “definite”, “specific”, “particular”, “individualized” etc. than *a car*, but, as noted above, *a car* certainly denotes a particular or specific car as far as the speaker is concerned. The difference is that the reference of *the car* in (2) is assumed to be clear to the hearer as well as the speaker. This is the first crucial insight; whereas in the case of an indefinite noun phrase the speaker may be aware of what is being referred to and the hearer probably

¹ We will see, however, that articles can encode more than definiteness or indefiniteness, and that they have been argued to have a quite different principal function, at least in some languages.

not, with a definite noun phrase this awareness is signalled as being shared by both participants. One would typically utter (1) where the car in question has no place yet in the hearer's experience, and is being newly introduced to it. (2) would be used where the hearer knows or has seen the speaker's new car. She may be at the wheel right now, or they may be standing looking at it together in her drive; or it may be that the hearer has not yet seen the car in the speaker's possession, but was aware that she had been looking over a particular car in a showroom recently.²

Examples like these have led to a view of definiteness known as the **familiarity hypothesis**. *The* signals that the entity denoted by the noun phrase is familiar to both speaker and hearer, and *a* is used where the speaker does not want to signal such shared familiarity. The familiarity hypothesis has a long history, and its first full presentation is in Christophersen (1939), a work which has greatly influenced much subsequent writing on the subject. The major recent work in this tradition is Hawkins (1978), and the discussion I give here owes much to this account.

As further illustration, consider (3)–(12):

- (3) Just give **the shelf** a quick wipe, will you, before I put this vase on it.
- (4) Put these clean towels in **the bathroom** please.
- (5) I hear **the prime minister** behaved outrageously again today.
- (6) **The moon** was very bright last night.
- (7) An elegant, dark-haired woman, a well-dressed man with dark glasses, and two children entered the compartment. I immediately recognized **the woman**. **The children** also looked vaguely familiar.
- (8) I had to get a taxi from the station. On the way **the driver** told me there was a bus strike.
- (9) They've just got in from New York. **The plane** was five hours late.
- (10) **The president of Ghana** is visiting tomorrow.
- (11) **The bloke Ann went out with last night** phoned a minute ago.
- (12) a. **The fact that you've known them for years** is no excuse.
 b. We were rather worried by **the prospect of having to cook for six for two weeks**.

² This is something of a simplification, skirting a number of issues subject to debate. First, there is dispute over whether definite noun phrases can be referring expressions, or whether it is rather speakers who sometimes refer using them. Second, if definites can refer or be used to refer, it is less clear that reference is involved in the case of indefinites like *a car* here. Nevertheless, the distinction drawn captures a clear intuition, and its expression in terms of speaker's and hearer's familiarity with a referent is standard in at least the less technical literature. I return to this in Chapter 4.

Examples (3)–(5) show **situational** uses of *the*, in that the physical situation in which the speaker and hearer are located contributes to the familiarity of the referent of the definite noun phrase. In (3) the situation is the immediate, visible one; the shelf is familiar to speaker and hearer in that it is before their eyes. In (4) the situation is still relatively immediate, though the referent of the definite noun phrase is probably not visible; in a particular house, the hearer would most naturally take it that the reference is to the bathroom of that house. In (5) the relevant situation is wider; in a particular country, the reference to *the prime minister* would normally be taken to be to the prime minister of that country; the individual concerned is not personally known to the hearer, but is familiar in the sense of being known to exist and probably known by report. (6) can be regarded as a situational use in which the situation is the whole world, or as a use in which familiarity stems from **general knowledge**. Thus *the moon* is taken to refer to the particular moon associated with this planet, or to a unique entity forming part of the hearer's general knowledge.

In (7) we have examples of **anaphoric** *the*. The referents of *the woman* and *the children* are familiar not from the physical situation but from the linguistic context; they have been mentioned before. In this example the previous mention takes place in an earlier sentence uttered by the same speaker, but it could equally well occur in part of the discourse spoken by another person, as in the following exchange:

- (13) A: An old man, two women and several children were already there when I arrived.
B: Did you recognize **the old man**?

It is significant that in (7) and (13) the earlier mentions of the woman, the children and the old man take the form of indefinite noun phrases; new referents are introduced into the discourse in this form because they are so far unfamiliar to the hearer.

Examples (8) and (9) are **bridging cross-reference** or **associative** uses, and can be thought of as a combination of the anaphoric and general knowledge types. In (8) the driver has not been mentioned before, but there has been mention of a taxi, and it is part of our general knowledge that taxis have drivers. The idea is that the mention of a taxi conjures up for the hearer all the things that are associated with taxis (a driver, wheels, seats, the fare etc.), and any of these things can then be referred to by means of a definite noun phrase. So the referent of *the driver* is familiar through association with the antecedent *a taxi*. (9) is particularly interesting because the antecedent which warrants the definite *the plane* is not even a noun phrase. But travelling from New York to most places necessarily involves some form of conveyance, with an aircraft being the most likely if the present conversation is taking place in, say, Manchester.

There are two possible ways of characterizing (10). The hearer may not know the president of Ghana personally, nor even have heard of him, but will know from his knowledge of the world that there probably is such an individual. The alternative characterization involves taking the definite article to be modifying not *president of Ghana*, but just *president*, so that *of Ghana* is a phrase added to provide clarifying information and not itself within the scope of *the*. On this view, the prepositional phrase has the same function as the previous mention of *a taxi* in (8): to provide a trigger for the association that familiarizes the definite noun phrase. If this is correct, then it is possible for an associative use of *the* to be based on following as well as preceding information. A similar treatment seems appropriate for (11). In this example, the familiarity of *the bloke* depends on the following relative clause. Assume that the hearer did not even know that Ann had gone out last night. The relative clause informs him of this, and also informs him that she went out with someone. The familiarity of *the bloke* then consists of its association with this succeeding information.

Finally, consider (12). Here, *that you've known them for years* is the fact in question, and *having to cook for six for two weeks* is the prospect. So these clauses, again following rather than preceding the definite noun phrase, act as “antecedent” for *the fact* and *the prospect*, which are therefore anticipatory anaphoric (or “cataphoric”) uses.

It may be already clear from this presentation that the concept of familiarity as an explanation for the definite–indefinite distinction is not unproblematic. It is fairly straightforward for examples like (3)–(6), where the hearer is genuinely acquainted with the referent, (7), where previous mention makes the referent familiar (by report rather than direct acquaintance), and even (8), where the fact that taxis always have drivers affords the same sort of familiarity as in (7). But getting from New York to Manchester does not necessarily involve flying; the association appealed to in (9) is certainly real, but can one really say that the plane was in any sense known to the hearer before the utterance of the second sentence of this example? In (10), the hearer would normally be prepared to accept that Ghana has a president, but that is not the same as knowing this person. In (12), where cataphoric information is appealed to, one can claim that the necessary familiarity is established after the utterance of the definite noun phrase, but in (11), the fact that Ann went out with a man is not expressed in the relative clause; in fact, *the bloke* can be replaced by *a bloke*, without changing the referent, which seems to make it clear that the information in the relative clause is not such as to establish the familiarity that would make *the* obligatory.

Because of considerations like these, many linguists basically sympathetic to the familiarity thesis prefer to see definiteness as being about **identifiability**. The idea is that the use of the definite article directs the hearer to the referent of the

noun phrase by signalling that he is in a position to identify it.³ This view of definiteness does not altogether reject familiarity. Rather, familiarity, where it is present, is what enables the hearer to identify the referent. In such cases the hearer is invited to match the referent of the definite noun phrase with some real-world entity which he knows to exist because he can see it, has heard of it, or infers its existence from something else he has heard.

In the examples discussed above where familiarity seems rather forced, it is generally the case that the definiteness of the noun phrase confirms an association which is only probable or possible rather than known. In (9), the journey mentioned makes the involvement of an aircraft likely, and then the definite noun phrase *the plane* authorizes the hearer to associate its referent with this journey, confirming the possible association. It does this by indicating that its referent can be identified by the hearer, and the most straightforward identification is with a plane the travellers probably came on from New York. A similar association is involved in (10); Ghana probably has a president, and it is with this probable individual that the reference of *the president* is identified. But in this example, the phrase which provides the probable referent occurs after the definite noun phrase and is attached to it in such a way as to make the association certain rather than probable. In (11) the relative clause provides a context in which a referent for *the bloke* can be found. Ann went out last night with someone, and the referent of *the bloke* is that someone, even though the relative does not provide any information about the person (that it was a man, for example).

So while on the familiarity account *the* tells the hearer that he knows which, on the identifiability account it tells him that he knows or can work out which. Let us now consider a case where an explanation in terms of familiarity would be impossible. Back in the sitting-room which was the setting for (3), Ann is trying to put up a picture on the wall, and, without turning round, says to Joe who has just entered:

(14) Pass me **the hammer**, will you?

Joe looks around and, sure enough, sees a hammer on a chair. The difference between (14) and (3) is that, whereas the hearer in (3) knows there is a shelf in the room which provides an obvious referent for the definite noun phrase, Joe does not know at the time of Ann's utterance that there is a hammer in the room. He has to look

³ Note that the article itself does not identify the referent; *the* is a "grammatical word" with no descriptive lexical content, and therefore contains nothing which can itself identify a referent. The most it can do is invite the hearer to exploit clues in the linguistic or extralinguistic context to establish the identity of the referent. The article has been said by many writers to "pick out" an entity, but this is inaccurate; *the* may be about identifiability, but not identification.

for a referent, guided by the description *hammer*.⁴ The definite article tells Joe that he can identify the hammer Ann is talking about, and the verb *pass* (which tends to take things immediately available as complement, by contrast with *fetch*, *get*, *buy*) makes it almost certain that he will find it in the room. The referent of the definite noun phrase is unfamiliar to the hearer, but he is able to find a referent for it.

1.1.3 Uniqueness and inclusiveness

Identifiability certainly offers a more comprehensive picture than does familiarity, but there are also cases of definites for which an account in terms of identifiability is either not fully convincing or simply inadequate.

Associative uses of the definite article in general are problematic for identifiability; consider the following example:

- (15) I've just been to a wedding. **The bride** wore blue.

The definite reference *the bride* in (15) is successful because the hearer knows that weddings involve brides, and makes the natural inference that the reference is to the bride at the particular wedding just mentioned. But is it accurate to say that the hearer identifies the referent in any real sense? He still does not know who she is or anything about her. If asked later who got married that morning he would be in no position to say on the basis of (15), and if he passes the newly-wed in the street the next day he will not recognize her as the person referred to.

Many situational uses are also associative; they work because the hearer is able to associate a definite noun phrase with some entity which he expects to find in or associates with the situation. This is the case with the following:

- (16) [Nurse entering operating theatre]
I wonder who **the anaesthetist** is today.

A definite is possible because we take it for granted that operations involve anaesthetists. But it is clear from what is said in (16) that the speaker cannot identify the referent of the definite noun phrase, and does not necessarily expect the hearer to be able to. Both participants know there is or will be such an individual, but that is not identification. The point becomes all the clearer if we replace the definite article in (16) by a demonstrative:

- (17) I wonder who **that anaesthetist** is.

⁴ In the semantics literature, the term “description” is used of all material that ascribes properties to entities – including nouns as well as, more obviously, adjectives. A particularly important use of the word, especially in the philosophical literature, is in the term **definite description**, meaning an expression which ascribes a property or properties to a particular entity – in other words, a definite noun phrase. *The hammer*, then, is a definite description.

Again the speaker does not know the identity of the person referred to, but she is referring to a particular individual and expects the hearer to be able to pick out precisely which individual she means. This is not the case with (16), indicating that while demonstratives may require identifiability, definites do not.

Consider also a cataphoric case, where the definite article is sanctioned by a relative clause following the noun:

- (18) Mary's gone for a spin in **the car she just bought**.

In (18) the relative tells the hearer something about the car (the fact that Mary just bought it), but it does not help him identify it. He still would not know the car in question if he saw it (unless Mary was driving it).

What can be claimed about all these examples is that they involve the idea of **uniqueness**: the definite article signals that there is just one entity satisfying the description used. This uniqueness is generally not absolute, but is to be understood relative to a particular context. Thus in (15) there is just one bride at the wedding which triggers the association. In (16) the assumption is that there is just one anaesthetist taking part in the operation about to begin, but who it is is not known. And in (20) *the* conveys that Mary bought one car.

In the associative examples an indefinite article would seem unnatural, for various reasons; in (15), for example, the general knowledge on which the association is based includes an assumed normal pattern of one bride per wedding. But in (18) it is perfectly possible to substitute *a* for *the*:

- (19) Mary's gone for a spin in **a car she just bought**.

The most natural interpretation is still that only one car is involved, but the possibility is left open that Mary may have just bought more than one car. So the indefinite article does not signal non-uniqueness; rather it does not signal uniqueness. Indefinites are neutral with respect to uniqueness (though this will be qualified below).

As observed, the uniqueness of the definite article is usually relative to a particular context, but it can be absolute. This is the case with nouns which are inherently unique, denoting something of which there is only one. We can speak of *the sun* and *the universe*, but not normally of *a sun* or *a universe*; the qualification is important, because although for most purposes we think of our sun and our universe as the only entities to which those names apply, there are situations in which we might speak of our sun as one of many or entertain the possibility of there existing another universe. Nouns like *Pope* are also often thought of as inherent uniques, because there is usually only one at any given time; but of course if one looks across history there have been many Popes, and with this perspective it is reasonable to speak of *a Pope*. The fact that one can always find a context in which

a noun ceases to be uniquely denoting does not invalidate the point. Just as it is possible to claim that the count–mass distinction is basically valid despite the possibility of recategorizing any noun (as in the count use of the basically mass *milk* in *He strode up to the bar and ordered three milks*), so there is a class of inherently unique nouns. And such nouns, used as uniquely denoting, require the definite article.

Consider also the following immediate situation definite:

- (20) Beware of **the dog**.

This is intended to inform the reader that there is a dog in the vicinity, and that he is likely to meet it if he waits long enough or proceeds any further. One could argue that identifiability is involved, in that if he sees a dog nearby he is likely to connect it with the one mentioned in the notice. But there is no expectation that he will seek a referent for *the dog*; rather, (20) is equivalent to *There is a dog*. Uniqueness, on the other hand, does seem to offer an adequate account here, since an intrepid intruder could reasonably claim to have been misled if he found he had to deal with two dogs.

The uniqueness criterion is particularly attractive in cases where the referent is hypothetical, potential, or in the future:

- (21) **The winner of this competition** will get a week in the Bahamas for two.
 (22) **The man who comes with me** will not regret it.

Assuming the competition in (21) is not yet over and no one has yet agreed to accompany the speaker in (22), the winner and the man are certainly not yet identifiable. But they are unique, in that a single winner and a single male companion are clearly implied.

Finally, there are certain other modifying constituents of the noun phrase which are incompatible with the indefinite article; among these are superlatives, *first*, *same*, *only* and *next*:

- (23) Janet is **the/(*)a cleverest child** in the class.
 (24) You are **the/(*)a first visitor** to our new house.
 (25) I've got **the/(*)a same problem** as you.
 (26) He is **the/(*)an only student** who dislikes phonology.
 (27) I offered a discount to **the/(*)a next customer**.

Uniqueness offers an explanation for these facts, according to Hawkins (1978), since the unacceptability of the indefinite article seems likely to stem from a semantic incompatibility between an element of uniqueness in the meaning of the modifier

and the non-uniqueness of *a*. For although I have said that the indefinite article is neutral with respect to uniqueness, there are cases where choosing *a* rather than *the* implies non-uniqueness; this is a point I will return to. For the moment it will suffice to look at it in this way: if the descriptive material in the noun phrase indicates that the referent is unique, then the only appropriate article is the one that encodes uniqueness. This is the case with inherently unique nouns, and noun phrases containing superlatives etc. *Cleverest* means ‘cleverer than all the others’, and *first* means ‘before all the others’; so uniqueness can be argued to be involved here, as it obviously is with *only*. In (25), if the hearer has a single problem, or a single salient problem, as seems to be implied, then the speaker can have only one problem which is the same as the hearer’s. *Next* means ‘immediately following’, and given that customers are generally dealt with one by one, there can be only one customer who immediately follows the preceding one.

All the examples so far considered in this section have involved count nouns in the singular. But the definite article can occur equally well with plural count nouns and mass nouns, and the obvious question is: How can a definite noun phrase which is plural or mass have a referent which is unique (in the context)? The noun phrases *the pens* and *the butter* (the latter occurring with its usual mass value and not recategorized as count) cannot refer to just one pen and just one butter. Let us look at examples corresponding to those examined above, but with plural (the (a) sentences) and mass (the (b) ones) definite noun phrases:

- (28) a. We’ve just been to see John race. The Queen gave out **the prizes**.
b. We went to the local pub this lunch time. They’ve started chilling **the beer**.
- (29) a. [Nurse about to enter operating theatre]
I wonder who **the anaesthetists** are.
b. [Examining restaurant menu]
I wonder what **the pâté** is like.
- (30) a. We’re looking for **the vandals** who broke into the office yesterday.
b. I can’t find **the shampoo** I put here this morning.
- (31) a. Beware of **the dogs**.
b. Beware of **the electrified wire**.
- (32) a. We’re offering several prizes, and **the winners** will be invited to London for the presentation.
b. Fred’s decided to take up home brewing. He plans to sell **the beer** to his friends.
- (33) a. Janet and John are **the cleverest children** in the class.
b. This is **the best muesli** I’ve ever tasted.
- (34) a. You are **the first visitors** to our new house.
b. This is **the first rain** to be seen here for five months.

- (35) a. I've got **the same problems** as you.
b. All the family used to take their bath in **the same water**.
- (36) a. They are **the only students** who dislike phonology.
b. This is **the only water** you're likely to see for miles.
- (37) a. I offered a discount to **the next three customers**.
b. **The next water** is beyond those hills.

As a first attempt at a solution, one might propose that uniqueness still applies, but to sets and masses rather than to individuals. Thus the set or mass referred to by a definite noun phrase is the only set or mass in the context satisfying the description. But this does not work. In (28a), suppose there are three prizes. These form a set, but there is also the set consisting of the second and third prizes, that consisting of the first and second prizes, and so on; these, of course, are subsets of the set of three – and this is the point. Our intuition about (28a) is that the Queen gave out all the prizes, not some subset of the total; similarly in (28b), we assume that all the beer at this pub is now served chilled.

This points us to the proposal that definiteness, at least with plural and mass noun phrases, involves not uniqueness but **inclusiveness** (a term due to Hawkins (1978)). What this means is that the reference is to the totality of the objects or mass in the context which satisfy the description. So in the (a) examples of (29)–(32) the reference would be taken to be to all the anaesthetists about to take part in the operation, all the vandals involved in the break-in, all the dogs guarding the property, and all the winners in the competition. In the (b) examples, it is to all the pâté on offer in the restaurant, all the shampoo left there, the electrified wire surrounding the property as a whole, and all the beer Fred brews. In (33), Janet and John are the only children in the class meriting the description *cleverest*, and the muesli praised is the totality of the muesli in the speaker's experience deserving to be called the best. I leave the reader to work out how inclusiveness accounts for (34)–(37).

It appears, then, that with plural and mass nouns *the* is a universal quantifier, similar in meaning to *all*. As support for this position, consider the following:

- (38) a. I've washed the dishes.
b. I've washed all the dishes.
- (39) No you haven't, you've only washed some of them.

Our intuition is that (38a) and (38b) are equally false if there are still some dishes unwashed, and in that case (39) would be a reasonable retort to either. So *the* (in some uses) and *all* are very close in meaning, and the difference between them may be that *all* is simply more emphatic. But it seems unsatisfactory to say that *the* signals uniqueness with singular noun phrases and inclusiveness with plural

and mass noun phrases. In fact uniqueness can be assimilated to inclusiveness. When the noun phrase is singular, inclusiveness turns out to be the same as uniqueness, because the totality of the objects satisfying the description is just one. For example, the speaker in (16) is assuming there will be only one anaesthetist; so the total number of anaesthetists assumed to be involved is one.⁵

1.1.4 Identifiability, inclusiveness and indefinites

The relationship of the indefinite article to identifiability and inclusiveness is rather complex. We saw in relation to (18)–(19) that *a* is neutral with respect to uniqueness rather than signalling non-uniqueness. Where *the* is used the referent has to be unique: Mary bought one car. *A* allows this same interpretation, while also permitting an interpretation in which the car referred to is one of several. This picture is reinforced by sentences like the following:

- (40) I went to the surgery this afternoon and saw **a doctor**.

The doctor the speaker saw may have been one of several in the surgery, but not necessarily; (40) is perfectly compatible with there having been only one doctor there.

But there are other cases where *a* signals non-uniqueness and the choice of *a* rather than *the* makes a significant difference:

- (41) Pass me **a hammer**.

- (42) Janet ran well and won **a prize**.

These sentences clearly imply that there is more than one hammer in the situation to choose from, and Janet won one of a number of prizes.

It appears that, while *the* logically entails uniqueness with singular noun phrases, *a* is logically neutral with respect to this. But it carries a weaker implication of non-uniqueness. How this can be formalized will be discussed more fully in Chapter 7, but for the moment the point is that *a* may imply that the referent is non-unique, but this implication may be overridden. This can be illustrated in relation to (42), where it is possible to add material indicating that the referent is, after all, unique:

- (43) Janet ran well and won **a prize** – the only prize in fact.

⁵ It should be noted, however, that with singulars *the* is not (near-)synonymous with *all*. We use *the table* to refer to the only table in a certain context, but *all the table* denotes every part of the table, and **all table* is not a well-formed noun phrase.

There have been various reformulations of the basic insight behind inclusiveness, for example as “maximality” (Sharvy 1980, Kadmon 1987): the reference of a definite description is to the maximal set satisfying the description.

The fact that there is no contradiction in this sentence makes it clear that although in (42) the indefinite *a prize* strongly implies non-uniqueness, this implication is less central to the meaning of *a* than is uniqueness to that of *the*: (44) is much less acceptable than (43):

(44) ?Janet ran well and won **the prize** – one of several in fact.

We can summarize by saying that when a referent is inclusive in the context, *the* is normally used rather than *a*, because *a* implies non-inclusiveness (equivalent to non-uniqueness, since *a* only occurs with singulars). This implication can be overridden, however, making it clear that non-inclusiveness is not an entailment of the indefinite article. When the referent is not inclusive in its context, *a* must be used and *the* may not be, as (44) makes clear.

But there is a further use of *a*. This is when the entity referred to is not associated with what I have been calling a “context” – a physical situation or the previous discourse. First-mention uses exemplify this:

(45) I met a lion-tamer this morning.

Where there is no contextual set (other than perhaps the whole world) within which inclusiveness may or may not apply, *a* is used. And in such cases it seems to be a matter of non-identifiability rather than non-inclusiveness. The referent is taken to be unfamiliar to the hearer because it has not been mentioned before. This is probably also the best explanation for the indefinite in (19), where the referent may well be unique. This use is probably not to be seen as cataphoric (the relative postmodifier establishing a domain for uniqueness or familiarity); rather, the entire noun phrase, including head noun and relative, is treated as non-identifiable, and is therefore indefinite.

1.1.5 A unified account?

We have seen that familiarity can be subsumed under identifiability, and that uniqueness is merely a special case of inclusiveness, resulting from the singularity of the noun phrase. So the question now is: Can we make a choice between identifiability and inclusiveness? Is one of them right and the other wrong, or are there two kinds of definiteness?

In section 1.1.3 we examined a number of uses of *the* which can be accounted for by inclusiveness but not (or not very convincingly) by identifiability. Indeed the inclusiveness account could be extended to many other examples that are not problematic for identifiability. If certain uses can be handled by inclusiveness and not by identifiability, and other uses can be handled by either hypothesis, then inclusiveness must be preferred. Consider for example (14), repeated here:

- (14) Pass me the hammer, will you?

This example is unproblematic for the identifiability account. The hearer was not previously aware of the presence of a hammer in the room, but takes it from the definite noun phrase that there is one which he can identify; he looks around and finds the referent. But inclusiveness works equally well. If on looking around the hearer saw three hammers, he might have to ask back which one was meant; the definite reference would have failed, because the hearer could not identify the referent. Identification is thus only possible in (14) if the desired hammer is unique in the context. So both identifiability and inclusiveness work here as explanations for the appropriateness of *the*.

However, there are cases where identifiability works and inclusiveness does not. Consider the following:

- (46) [In a room with three doors, one of which is open]
Close **the door**, please.
- (47) [In a hallway where all four doors are closed. The speaker is dressed in coat and hat, and has a suitcase in each hand]
Open **the door** for me, please.
- (48) [Ann, fixing her motorbike, is examining a large nut. Behind her, just out of reach, are three spanners, two of them obviously far too small for the nut]
Pass me **the spanner**, will you?
- (49) [Two academics]
A: How did the seminar go?
B: Fine. **The student** gave an excellent presentation, which generated a really good discussion, with all the other students contributing well.

In (46), an immediate situation use, the door referred to is not unique; but it is easily identified because of the verb – you can only close an open door. (47) is similar; the door is not unique, but the speaker's state of preparedness for a journey makes it obvious that the street-door is meant. In (48) the hearer is expected to be able to work out that only one of the three spanners can possibly fit the bill and identify that as the intended referent. And in (49), the nature of a seminar makes it clear that one student in the group stood out as having a special task, and this individual will be taken to be the referent of *the student*. All these examples represent perfectly normal uses of *the*, and they cannot be adequately accounted for by inclusiveness.

We have, then, some usage types which can only be accounted for by identifiability, some which can only be accounted for by inclusiveness, and some which both theories account for equally well. This is not a satisfying conclusion, and I shall return to the matter in Chapter 7. But for the moment let us settle for a view

of definiteness as involving either identifiability or inclusiveness, or both: if the reference of a noun phrase is characterized by either property, then that noun phrase should be definite. Bear in mind, however, that the two properties are in principle independent of one another, even if in many examples the presence of one follows from the presence of the other; the two theories make quite distinct claims.⁶

It is also possible that what we are calling definiteness is in fact two or more distinct semantic categories, which happen to have the same lexical or morphological realization in English. A way to test this possibility would be to look for languages in which different kinds of “definiteness” are expressed in different ways, by different articles for example. Or it may be that definiteness is a unified phenomenon, but that neither identifiability nor inclusiveness is the correct characterization. We will come back to these speculations, assuming for the present that definiteness is a single category.

1.2 Types of definite noun phrase

This section surveys, still informally, and still using mainly English data, the range of noun phrase types which have definiteness as part of their meaning. Including the definite article in a noun phrase is not the only way of making definite reference to some entity. There are several other kinds of noun phrase which appear either to express the inclusiveness of the referent or to indicate that the referent is identifiable. The following noun phrases all have much in common semantically with definite noun phrases containing *the*: *that man*, *these houses*; *Ann*, *Venice*; *my car*, *a friend's house*; *they*, *us*; *all writers*, *every shop*.

1.2.1 Testing for definiteness

Articles like *the* and *a* are part of the larger class of **determiners** – another term used here informally, without commitment to the grammatical category of the items concerned, to cover non-adjectival noun phrase modifiers such as *this*, *several*, *our*, *all*. Some of these are definite determiners, differing from *the* in that they combine definiteness with some other semantic content. The

⁶ A number of writers present definiteness in terms which suggest an attempt to combine or reconcile the two accounts, characterizing definites as “uniquely identifying”. One should not be misled by this; it either represents a failure to appreciate the difference between uniqueness and identifiability, or is merely equivalent to “identifying”, or “identifying unambiguously”. A good example is the account of definiteness given by Leech (1983: 90–3). Leech says that the use of *the* conveys that there is a referent that can be uniquely identified by speaker and hearer. He adds that “uniquely” means that “we should be able to select the one X concerned from all other X's”. There is nothing here about the referent being unique (in the context). But then he refers back to this characterization as “the uniqueness implicature associated with *the*”, which ensures that *the postcard I got from Helen last week* implies that there exists only one such postcard. This is by no means an isolated instance of confusion of identifiability and uniqueness.

combination may be a necessary one, because this other semantic content entails definiteness or is incompatible with indefiniteness. Or it may be that two semantic properties, in principle independent of one another, happen to be jointly encoded by a portmanteau morpheme. I shall follow the established practice of representing the grammatical and semantic content of lexical items and morphemes by features. So *the* is [+ Def] and *a* is [– Def]. Other determiners may be characterized as [+ Def] along with other feature specifications.⁷ I have adopted the term “simple definites/indefinites” to describe noun phrases with *the* or *a*; so other definite noun phrases, in which [+ Def] is present as a consequence of, or otherwise in combination with, another feature, are **complex definites**.

It has long been recognized that a number of syntactic environments either do not admit or admit only with some difficulty a definite, or conversely an indefinite, noun phrase. These environments have been used as diagnostics, to determine whether a given noun phrase type is definite or not. The following examples show how some of these “definiteness effects” distinguish between simple definites and simple indefinites; the constructions in (50)–(53) were first exploited by Postal (1970) and that in (53) is discussed in detail by Milsark (1979).

- (50) a. Big as the boy was, he couldn’t lift it.
 b. ?Big as a boy was, he couldn’t lift it.
- (51) a. The house is mine.
 b. ?A house is mine.
- (52) a. which/some/all of the women
 b. *which/some/all of (some) women
- (53) a. Is there a dictionary in the house?
 b. ?Is there the dictionary in the house?

These diagnostics are not foolproof, and not too much reliance should be placed on them. For example, the structure in (50) admits an indefinite noun phrase interpreted generically (as referring to an entire class), and similar counterexamples can be devised to some of the other tests:

⁷ Although writers are commonly noncommittal as to whether features are grammatical (morphosyntactic) or semantic, my discussion so far (with examination of usage types and the attempt to characterize definiteness in terms of notions like inclusiveness and identifiability) implies that I take the features I introduce to be semantic. Indeed nearly all discussion in the literature of definiteness and related concepts like demonstrativeness takes these to be semantic or pragmatic. I will in fact argue in later chapters that definiteness (though not necessarily demonstrativeness) is a grammatical category, not a semantic one (though it is related to a semantic/pragmatic concept). For the moment, I assume that [+ Def] is a defining characteristic of certain determiners, so that a noun phrase with one of these determiners is definite, even if neither inclusiveness nor identifiability seems to be involved in the interpretation. This assumption will be important in 1.2.4 below.

- (54) Big as a bus is, it can easily pass through this gap.
 (55) A house is mine if I pay for it.

Nevertheless, the examples (50)–(53) do draw a recognizable distinction, and these diagnostics offer a rough guide. If applied to a range of putative complex definites, they range these clearly with *the*:

- (56) a. Big as **that boy** was, he couldn't lift it.
 b. **That house** is mine.
 c. which/some/all of **those women**
 d. ?Is there **that dictionary** in the house?
 (57) a. Tall as **Nuala** is, she won't be able to reach it.
 b. **Fido** is mine.
 c. (*which)/some/all of **Paris**
 d. ?Is there **Peter** in the house?
 (58) a. Big as **my cousin** is, he can't lift it.
 b. Make yourself at home – **my house** is yours.
 c. which/some/all of **the students' essays**
 d. ?Is there **Rachel's racket** in here?
 (59) a. Clever as **you** are, I bet you won't solve it.
 b. **I** am yours.
 c. which/some/all of **us**
 d. *Is there **him** here yet?
 (60) a. Strong as **every contestant** is, they'll never shift it.
 b. **All hats** are yours.
 c. (?which)/(?some)/(?all) of **all the men**
 d. ?Is there **every visitor** here?

The partitive structure in the (c) examples shows some deviations from the pattern seen with *the*; *which of* cannot be followed by a proper noun because these are always singular, and noun phrases with *all* do not readily occur at all in partitives, for reasons which are not obvious. Otherwise the only point of note is that personal pronouns are even less acceptable in the *there is/are* construction than are other definites. In the following sections we will look more closely at demonstratives, proper nouns, possessives, personal pronouns, and determiners like *all* and *every*.

1.2.2 Demonstratives

Demonstratives are generally considered to be definite, but it is clear that their definiteness is not a matter of inclusiveness. A sentence like *Pass me that book* is likely to be used in a context where there is more than one possible referent corresponding to the description *book*, and the utterance may well be

accompanied by some gesture indicating which book the speaker has in mind. Hawkins (1978) argues that demonstrative reference always involves a contrast, clear or implied, between the actual referent and other potential referents. On the view so far adopted here, identifiability is what links demonstratives with the definite article. In uttering *Pass me that book*, or *I've read this book*, the speaker assumes that the hearer can determine which book is intended, by contrast with *Pass me a book* or *I've read a book* (the former not involving any intended referent).

But identifiability is only part of the semantic content of demonstratives. They are often grouped with the varied class of words which express **deixis**. Deixis is the property of certain expressions and categories (including tense and grammatical person) of relating things talked about to the spatio-temporal context, and in particular to contextual distinctions like that between the moment or place of utterance and other moments or places, or that between the speaker, the hearer, and others. Demonstratives like *this* and *that* are deictic because they locate the entity referred to relative to some reference point in the extralinguistic context. The contrast between *this/these* and *that/those* is to do with distance from the speaker; *this book* denotes something closer to the speaker than does *that book*. This distance is not necessarily spatial; it may be temporal (*that day* referring to some past or future occasion, as opposed to *this week*, meaning the present week), or emotional (*There's that awful man here again*, *What about this present you promised me?*). For discussion of these possibilities see Lakoff (1974). *This* and *that* are often termed **proximal** and **distal** demonstratives, respectively. But it is possible to relate this distance contrast to the category of person. *This* is used to refer to some entity which is close to or associated in some way with the speaker, or with a set of individuals which includes the speaker; so *this article* could be 'the article which I am reading', 'the article which you and I are discussing', or 'the article which you, I and they are interested in', among other possibilities. Now this concept of a set of one or more individuals which includes the speaker is the definition of "first person"; it corresponds to the pronouns *I* (set of one) and *we* (set of more than one). *That* is used where the referent is associated with a set including hearer but not speaker (second person) or a set including neither speaker nor hearer (third person):

(61) Show me **that** (?**this**) letter you have in your pocket.

(62) Tell her to bring **that** (?**this**) drill she has.

This is certainly possible in these examples, but it would imply that the letter or drill is in some way associated with the speaker (or a set including the speaker). Thus, for example, the letter may have already been the subject of discussion between speaker and hearer; and the speaker may have been previously thinking about the amazing drill she has recently heard one of her friends has acquired. But in the

absence of such factors (or of the speaker's desire to communicate such factors), *that* is the appropriate demonstrative: in (61) the letter is in the possession of the hearer (second person), and in (62) it is someone not present in the discourse situation (third person) who has the drill. So it would be reasonable to speak of *this* as a first-person demonstrative and *that* as a non-first-person demonstrative.

An important question is whether this deictic element is the defining characteristic of demonstratives, distinguishing them from the definite article. An affirmative answer is given by many writers. Sommerstein (1972) and J. Lyons (1975, 1977) assume that the distance component is the only thing that distinguishes *this* and *that* from *the* (so that *this* marks proximity, *that* marks distance or non-proximity, and *the* is neutral with respect to distance); and Anderson and Keenan (1985) comment that a demonstrative system with only one term "would be little different from a definite article". But there are good reasons for rejecting this view.

English *that* is sometimes neutral with respect to distance or person, as when used pronominally in relative constructions:⁸

(63) She prefers her biscuits to **those** I make.

(64) I want a coat like **that** described in the book.

Taking the deictic opposition to be expressed by a feature [\pm Prox], then either these demonstrative occurrences do not carry this feature, or the negative value [$-$ Prox] characterizing *that* and *those* can include neutrality with respect to distance. On the assumption that *the* too is either unmarked for [\pm Prox] or redundantly [$-$ Prox], there must be some other feature distinguishing *the* and *that*.

More striking evidence is afforded by languages which have a demonstrative unmarked for any deictic contrast which is either the only form in the system or is distinct in form from the terms which are deictically marked. Egyptian Arabic has basically a one-demonstrative system with no deictic contrast: *da* 'this' or 'that'

⁸ Stockwell, Schachter and Partee (1973) regard *that* and *those* in such sentences as forms of the definite article rather than demonstratives, because of a supposed complementary distribution between these forms and the usual simple definite form – *that* and *those* occurring with mass and plural value before a post-nominal modifier, and *the one* occurring with singular count nouns. They would disallow *the ones* for *those* in (63) and *that* (as opposed to *the one*) in (64), but they are simply mistaken in this. The claim of complementary distribution fails particularly clearly in the case of plurals like (63), where *the ones* and *those* are equally acceptable. Stockwell, Schachter and Partee do have a point as regards singulars like (64), for which the distribution they present does represent the general tendency; and there is no obviously simple definite alternative to *that* relating to a mass noun phrase (as in *I prefer this butter to that you got from the market*), unless we propose that the "free relative" *what* is this alternative (*I prefer this butter to what you got from the market*). But the fact that the complementary distribution claimed is only partial and optional indicates that an analysis in which the demonstrative forms really are demonstratives cannot be ruled out. But if we take the forms in (63)–(64) to be true demonstratives, it is evident that they do not express any degree of distance or association with person.

(Mitchell 1962). There are ways of indicating the relative distance of an object from the speaker, but such information is not lexicalized in the demonstrative system. But *da* is not a definite article; Egyptian Arabic has a definite article *?il*, distinct from the demonstrative. Another case is French, where distance is expressed by a suffix *-ci* or *-là* on the noun, so that the information usually thought of as belonging to the demonstrative is divided between two morphemes: *ce bateau-ci* ‘this boat’, *ce bateau-là* ‘that boat’. But this suffix can be freely omitted, so that no information about distance is conveyed: *ce bateau* ‘this/that boat’. But again, *ce* is still distinct, in form and meaning, from the definite article *le*.⁹

So deixis is a usual but not invariant property of demonstratives. There must be some other property, then, that distinguishes them from the definite article. Let us call it [\pm Dem].¹⁰ So *this/these* is [+ Def, + Dem, + Prox]. Views differ very widely, however, on what the [\pm Dem] distinction amounts to; put differently, writers who agree that demonstratives have some distinctive property apart from deixis are divided on what it is.¹¹ Hawkins (1978) claims that demonstratives are distinguished by a “matching constraint”, which instructs the hearer to match the referent with some identifiable object, that is, some object which is visible in the context or known on the basis of previous discourse. This constraint captures the observation that, whereas with a simple definite the referent may be inferrable on the basis of knowledge of the world (as in (65), cars being known to have engines), with a demonstrative the referent must be given in the linguistic or non-linguistic context – thus the impossibility of (66):

(65) I got into the car and turned on **the engine**.

(66) *I got into the car and turned on **this/that engine**.

Hawkins’s matching constraint looks rather like identifiability, though a more restricted notion of identifiability than is involved in definiteness. See also Maclaran

⁹ Harris (1977, 1980) argues that *ce* is rivalling *le* as definite article, the latter form tending to become an unmarked determiner with the function of carrying agreement features (number and gender), but this process is certainly not complete yet. The sentences *Passe-moi le marteau* ‘Pass me the hammer’ and *Passe-moi un marteau* ‘Pass me a hammer’, in immediate situation use, show the same contrast as the English glosses.

¹⁰ The conclusion that the deictic content of demonstratives that have it is not what distinguishes them from the definite article becomes inescapable in view of the fact, which will emerge in Chapter 2, that there are languages in which non-demonstrative definite articles show these deictic distinctions.

¹¹ It is important to be aware of a certain variation in the literature in the use of the term “deixis”. I have been using it here to denote the distinctions of proximity to the speaker or association with a particular person which make it possible to locate entities in the context of utterance relative to others. But the term is also often used to denote the basic property common to all demonstratives, the property expressed here by the feature [+ Dem]. On this use, “deictic” is equivalent to “demonstrative”, being applied to forms which “point out”.

(1982) for a modified version of this account. The position I adopt here is that a demonstrative signals that the identity of the referent is immediately accessible to the hearer, without the inferencing often involved in interpreting simple definites. This may be because the work of referent identification is being done for the hearer by the speaker, for example by pointing to the referent. The deictic feature typically expressed on a demonstrative plays a similar role to pointing, guiding the hearer's attention to the referent. This suggests a necessary connection between [+ Dem] and [+ Def], the former implying the latter. I take demonstratives, then, to be necessarily definite.

1.2.3 Proper nouns

The term “proper noun”, or “proper name”, is applied to a very heterogeneous set of expressions, including *John*, *The Arc de Triomphe*, *South Farm Road*, some of which have internal grammatical structure and contain descriptive elements, and some of which do not. I limit my attention here to names like *John* and *Paris*, which (though they have an etymology) are not generally thought of as having any descriptive semantic content, or as having any meaning independent of the entity they name. The name *Paris* is applied to a particular city, but tells us nothing about that city; of course we would normally expect the bearer of the name *John* to be male, but it is argued that this is not part of the meaning of the name. By contrast, the common noun *man*, in being applicable to a particular individual, is so in virtue of the fact that that individual satisfies certain descriptive criteria (being human, male, adult). Proper nouns are often said to be referring expressions but to have no sense. They are also sometimes said to be logically equivalent to definite descriptions, in being uniquely referring expressions.

There may be millions of people called *John* and there are several towns called *Paris*, so context is important for the identification of the referent, as with definite descriptions. But a common view is that we use proper nouns as if they were absolutely unique, corresponding more closely to inherently unique definites (like *the sun*) than to possibly contextually unique definites (like *the man*). When we are conscious of there being more than one possible referent for the name *John* we can either expand it to a fuller proper noun (*John Smith*) or recategorize it as a common noun and add some descriptive material (*the John I introduced you to last night*). I shall assume this to be correct as an outline of how proper nouns are used; for a detailed and clear discussion of the complexities of proper nouns see J. Lyons (1977: 177–229).

It is clear that the uniqueness of reference of proper nouns is what aligns them with definites, though it may be added that this very uniqueness will generally ensure the identifiability of their referent. But a number of questions arise at this point. How do proper nouns differ from inherently unique nouns like *sun*? They

have in common that they are both generally used as though they denote a unique entity, but they differ grammatically: *sun* behaves like a common noun in that it takes the article, or some other definite determiner (*the sun*, *that lucky old sun*); *John*, unless recategorized, generally does not, and in fact is not only a noun, but also a complete noun phrase. One answer is that nouns like *sun* denote singleton sets, while proper nouns denote individuals; this would be in keeping with the view that proper nouns have reference but not sense. Another, implying that proper nouns do have sense, is that both types of noun denote singleton sets, but in the case of *sun* the set just happens to have only one member, while the set satisfying *John* is by definition a single-member set.

This latter proposal goes some way towards answering the question why proper nouns in English do not take the definite article. If by definition they denote a singleton set, there is no need to signal the uniqueness of their referent. But are they then in fact definite, or merely semantically similar in some way to definite noun phrases? In other words, is the feature [+ Def] present in proper nouns, and if so, where? If proper nouns are [+ Def], this feature would appear to be on the noun, given the lack of a determiner. But it seems to be clearly a determiner feature in common noun phrases. It would be preferable to be able to say that the definiteness feature occurs in one place only, and in general the determiner seems the most probable locus (unless we say a grammatical feature can have its locus in a phrasal category, so that it is the noun phrase, not the noun or the determiner, which carries [\pm Def]). If we assume that the feature [+ Def] pertains only to determiners, it may be that proper nouns are accompanied by a phonetically null determiner, or that the feature does not after all appear on proper nouns. One proposal along these lines is Lyons (1995c), where it is argued that proper nouns in English are in fact indefinites, and their apparently definite behaviour comes from their being generics – generic noun phrases anyway (whether definite or indefinite) showing similar distributional behaviour to definite non-generics. This idea will be taken up again in Chapter 4.

In a number of languages, at least some types of proper noun (most typically personal names) do regularly take the definite article. Examples are Classical Greek (*ho Sōkratēs* ‘Socrates’, *hē Hellas* ‘Greece’), Catalan (*l’Eduard*, *la Maria*). It may be that these names should be treated as no different from uniques like English *the sun* – and thus unlike English proper nouns. It is clear that they are definite noun phrases, with [+ Def] encoded in the determiner.

1.2.4 Possessives

Under this heading I include determiners like *my*, *their* (together with their pronominal forms *mine*, *theirs*), and also “genitive” forms like *Fred’s*, *the woman’s*, *that man next door’s*. These genitives are clearly full noun phrases, and

the 's ending (which may be a genitive case morpheme or a postposition, among other possibilities) appears right at the end of the whole phrase. Possessives like *my* are also formed from noun phrases, since they are derived from personal pronouns. These possessive forms of noun phrases occur as modifying expressions within other noun phrases, as illustrated in (67), where each of the bracketed expressions is a noun phrase (or a derivative of a noun phrase):

- (67) a. [[my] cousin]
 b. [[Fred's] only friend]
 c. [[that man next door's] car]

Now, in English at least, possessives render the noun phrase which contains them definite, as shown by the diagnostics introduced in 1.2.1. And the phrases in (67) could be roughly paraphrased by the expressions *the son/daughter of my aunt and uncle*, *the only friend Fred has*, *the car belonging to that man next door* – clearly definite noun phrases, beginning indeed with the definite article.¹² The same applies to many other languages; as a further example, with a different word-order

¹² It might be supposed that what makes these noun phrases definite is that the noun phrase underlying the possessive expression (*me*, *Fred*, *that man next door* in (67)) is definite – the definiteness of the embedded noun phrase is somehow transferred to the matrix noun phrase. Precisely this explanation has been offered by a number of writers for the definiteness of the corresponding structures in Semitic languages; we will return to this in Chapter 3. This would mean that if an indefinite noun phrase (such as *a woman*) were made the basis of the possessive (thus *a woman's*), the matrix noun phrase (*a woman's drink* for example) would be indefinite. And this may seem to be the case, since such noun phrases occur fairly readily after existential *there is/are*:

- (i) There's a woman's drink on the shelf here.

But this impression is mistaken. Noun phrases like *a visitor's hat*, *a friend of mine's cousin*, containing an indefinite possessive, are definite, and are naturally paraphrased by the clearly definite *the hat of a visitor*, *the cousin of a friend of mine*. The impression of indefiniteness comes from the fact that examples like *a woman's drink* are structurally ambiguous, as indicated in (ii)–(iii), where the possessive phrase is bracketed:

- (ii) a [woman's] drink
 (iii) [a woman's] drink

The structure we are interested in here is the one in (iii), where the possessive is the indefinite *a woman's*. The matrix noun phrase is definite, natural paraphrases being *the drink belonging to a woman*, *the drink left behind by a woman*, etc. In (ii) the possessive expression *woman's* is probably not a full noun phrase, and is therefore not indefinite. The indefinite article *a* is not part of the possessive expression, but is a modifier of the matrix noun phrase and accounts for its indefiniteness. The sense of (ii) is something like 'a drink suitable for women'. For a detailed discussion of this distinction between "inner genitives" (as in (ii)) and "outer genitives" (as in (iii)) see Woisetschlaeger (1983). The observation stands, then, that in English a possessive noun phrase, whether itself definite or indefinite, renders its matrix noun phrase definite.

pattern, consider Irish, in which possessives derived from pronouns appear pre-nominally (*mo/do hata* ‘my/your hat’), and those based on “full noun phrases” are genitive case forms occurring to the right of the head noun (*hata an fhir* (hat the+GEN man+GEN) ‘the man’s hat’). In neither case can the head noun be modified by the definite article (**an mo hata*, **an hata an fhir*), yet the matrix noun phrase must be understood as definite; these examples are equivalent to ‘the hat belonging to me/you/the man’. But in many other languages possessives do not impose a definite interpretation on the matrix noun phrase. Italian has only pronoun-derived possessive determiners (*mio* ‘my’ etc.); possession with full noun phrases is expressed prepositionally (*il libro di Carlo* (the book of Carlo) ‘Carlo’s book’) – a structure which does not seem to impose definiteness in any language. Now the Italian translation of the definite *my book* is *il mio libro* (the my book), in which the definiteness is conveyed by the article *il*; this article may indeed be replaced by the indefinite article *un*, and the matrix is then indefinite: *un mio libro* (a my book) ‘a book of mine’. The point is that the presence of the possessive *mio* has no bearing on whether the matrix noun phrase is definite or indefinite. The situation is similar in Classical Greek, which uses genitive case forms for full noun phrase possessives. To express the definite ‘the man’s horse’, the possessive noun phrase ‘the man’s’ (itself definite and therefore having the definite article) is preceded by the article which renders the matrix noun phrase definite, with the result that two articles occur in sequence: *ho tou andros hippos* (the+NOM the+GEN man+GEN horse+NOM).

This difference between English and Irish on the one hand and Italian and Greek on the other is discussed by Lyons (1985, 1986a), in terms of a typological distinction between “DG languages” (English, Irish) and “AG languages” (Italian, Greek).¹³ The difference between the two types is claimed to reside in the structural position occupied by the possessive; what it amounts to essentially is that in DG languages possessives appear in a position reserved for the definite article and other definite determiners, but in AG languages they are in adjectival or some other position. This claim is controversial, and I shall return to a more detailed discussion of it. For the moment we can say that in some languages a possessive induces definiteness in the matrix noun phrase while in other languages it does not. The traditional assumption that possessives are definite determiners, stated without further comment in many descriptive grammars and in much recent theoretical work – presumably because possession is assumed to entail definiteness – is misguided. It reflects a lack of awareness of the AG phenomenon. We have seen above that

¹³ “DG” and “AG” stand for “determiner-genitive” and “adjectival-genitive”, respectively (though it is not claimed that possessives are necessarily determiners in the first type and adjectives in the second).

[+ Dem] is always accompanied by [+ Def], apparently because demonstrativeness is semantically incompatible with indefiniteness. But this is not the case with possession.¹⁴ We shall return to the “definite constraint” on possessives in DG languages in Chapters 3 and 8.

There is one circumstance in which this definite constraint can be suspended, at least in some DG languages; this is when the noun phrase is in predicative position:

(68) Mary is **Ann’s friend**.

(69) I was once **Professor Laserbeam’s student**.

The noun phrases indicated may be understood as either definite or indefinite. There is not necessarily any implication that Ann has only one friend and Professor Laserbeam has only had one student; rather, the bold phrases in (68) and (69) are likely to be interpreted as equivalent to *a friend of Ann* and *a student of Professor Laserbeam*. But notice that this same interpretation is possible even where the predicative noun phrase with possessive modifier is marked as definite by the article: *I was once the student of Professor Laserbeam*.¹⁵

There is also a use of non-predicative noun phrases with possessive modification which cannot be characterized as either inclusive or identifiable:

¹⁴ This is easily demonstrated, even limiting the discussion to DG data. Note first that possessives do not necessarily express possession in the sense of ownership. The phrases *John’s house*, *John’s club*, *John’s annoyance*, *John’s son* may mean ‘the house John lives in’, ‘the club John is a member of’, ‘the annoyance John feels’, ‘the boy John has fathered’. Possessives express merely that there is a relationship of some kind between two entities (represented by the possessive noun phrase and the “possessum” – the matrix noun phrase). The nature of that relationship may be dictated lexically by the head noun, if it is relational or a body part for example; or it may be determined pragmatically, the context and background determining whether *John’s team* is the team John owns, the team John supports, or the team John plays in. Now, assume for the sake of argument that *John’s car* is ‘the car belonging to John’ and *John’s team* is ‘the team John supports’. If “possession” is semantically incompatible with indefiniteness in the possessum, the phrases *a car belonging to John* and *a team John supports* should be semantically anomalous. But they are not. In fact English has a possessive construction which can be indefinite, the “postposed possessive”: *a car of John’s*, *a team of John’s*. So the fact that a pre-head possessive in English imposes a definite interpretation on the matrix noun phrase has nothing to do with the semantics of possession.

¹⁵ This phenomenon is probably to be distinguished from the ambiguity as regards definiteness of predicative possessives with no overt possessum. In *That pen is mine*, *mine* is either merely the predicative form of *my* or a pronoun. If it is the latter, it is a definite noun phrase (‘That pen is the one belonging to me’) because the definite constraint applies to possessive pronouns. If it is the former, then the definite constraint does not apply, because the possessive is not part of a noun phrase (‘That pen belongs to me’). In Spanish these two readings would be clearly distinguished, and indeed the definiteness of the pronoun clearly marked by the article: *Esa pluma es mía* (predicative non-pronominal possessive), *Esa pluma es la mía* (pronominal possessive).

- (70) I'm going to stay with **my brother** for a few days.
(71) Joe has broken **his leg**.
(72) Oh, I've just torn **my sleeve** on that bramble.

There is no implication in these sentences that the speaker has only one brother and Joe only one leg, or that the hearer is in a position to tell which of the speaker's sleeves has been torn. In all these cases, an indefinite (*a brother of mine*) or a partitive (*one of my sleeves*) could have been used, but the definite structure is usually preferred. The referent is in each case one of a small number (people have only two legs or sleeves, and generally a limited number of brothers), and it is perhaps unimportant which of this number is at issue; nevertheless, neither inclusiveness nor identifiability applies. Since true indefinites can be used, there is no reason to consider the bold noun phrases in (70)–(72) to be indefinite. It seems rather that what we are dealing with here is definite noun phrases used where the conditions for definiteness do not strictly hold. This use is limited to noun phrases where the head noun is one of a small lexical class denoting mainly body parts, clothing, and family and other personal relationships – thus, inalienable and other intimate possession. It is not at all clear what lies behind these facts, but what makes them interesting is that they are not specific to English. Compare the following French example:

- (73) Jacques s'est cassé la jambe.
'Jacques has broken his leg.'

In fact this example does not contain a possessive, French generally using the definite article and a dative pronominal form to express possession with body parts (thus literally 'Jacques has broken to himself the leg'). The point is made all the more clearly here, given the definite article, that the "possessed" body part, though not unique or identifiable, is referred to by a definite noun phrase.

1.2.5 *Personal pronouns*

The personal pronouns are traditionally so called because they express grammatical person, but they have also long been recognized as definite and are often referred to as "definite pronouns" (by contrast with indefinite pronouns like *one* and *someone*). Postal (1970) proposes to account for the definiteness of personal pronouns by deriving them transformationally from definite articles, and I will adopt the essence of this account, though with some modifications first proposed in C. Lyons (1977).

Such an account has obvious attractions. In general English pronouns are part of determiner–pronoun pairs; so *this*, *that*, *one*, *some*, *all*, *each*, *several* and numerous other items can occur both pre-nominally (as determiners) and independently

(as pronouns). Two obvious exceptions to this are the definite article, which only occurs as a determiner (*Pass me the book*, but **Pass me the*), and third-person personal pronouns, which are always pronominal (*Pass me it*, but **Pass me it book*). A neat solution to this oddity would be to pair these together, and say that *he*, *she*, *it* and *they* are the pronominal correlatives of the determiner *the* (which is therefore third person). This idea is strongly supported by the observation that the first- and second-person plural pronouns *we* and *you* also occur pre-nominally:

- (74) a. **We Europeans** are experiencing some strange weather patterns.
 b. I don't trust **you politicians** an inch.

So *we* and *you* can be regarded as forms of the definite article, differing from *the* only in respect to person, but occurring both as determiners and as pronouns; suppletive variation is limited to the third person. In the singular, *you* does not occur freely as a determiner (**You linguist are relying on some pretty odd data*), but it is so used in exclamations: *You idiot!*, *You lucky bastard!*. The first-person singular pronoun does not occur at all in English pre-nominally (**I wish you'd leave me foreigner alone*; **I idiot!*). I can suggest no reason for this restriction, but it is, at least in part, a language-specific fact; German permits both the exclamatory use and the more general use with *ich* 'I':

- (75) Ich Esel!
 I donkey
 'Silly me!'
 (76) Ich Vogelfänger bin bekannt bei Alt und Jung . . .
 I birdcatcher am known to old and young . . .

Postal derives pronouns from an underlying full noun phrase consisting of an article plus a minimal noun head (with the form *one* in English). The latter is generally deleted, but surfaces in the third person when the noun phrase contains a restrictive modifier:

- (77) a. I met the one who/that Lucille divorced.
 b. I know the one with brown hair.
 c. I bought the green one.

The fact that normal third-person pronouns are (at least in colloquial standard) excluded with such modifiers (*?him who/that Lucille divorced*, *?her with brown hair*, **it green*), is further support for the analysis. The hypothetical noun *one* also shows up with first- and second-person articles in the presence of a preposed modifier: *us clever ones*.

Contesting this account, Sommerstein (1972) points out that while *we* and *you* can be stressed or unstressed (with a reduced vowel) when used pronominally (thus

[wi:] and [wɪ], [ju:] and [jə]), when they are used as determiners they are always stressed and with a full vowel, except in the exclamatory use. Thus the reduced form [jə] may occur in *You fool!*, but not in (74b). This ties in with the point made by Hawkins (1974) that pre-nominal *we* and *you* have as their third-person correlative, not *the*, but demonstratives, because of distributions like the following:¹⁶

- (78) a. the strongest/only soldiers in the army
 b. *we/you/these strongest/only soldiers in the army

The point that personal pronouns often have more in common with demonstratives than with *the* is taken up by C. Lyons (1977), who argues that Postal is basically right but that the English personal pronouns are forms both of the definite article and of the [– Prox] demonstrative. This claim is based on a case of complementary distribution between demonstratives and third-person pronouns.

The various uses and forms of English demonstrative pronouns must first be distinguished. There are two kinds of anaphoric use, **strict anaphora** as in (79), where the demonstrative is coreferential with a previously occurring noun phrase, and **identity of sense anaphora** as in (80), where there is no coreference but the previously occurring noun phrase supplies the understood descriptive content of the demonstrative.

- (79) Jim finally found a flat with a view over the park, and bought **that (one)**.
(80) I could never follow our first coursebook, and I'm glad we now use **this one**.

Then there is the non-anaphoric use, typically indicating something in the immediate situation. In fact the identity of sense example (80) is also situational, in that the demonstrative gets its reference from the immediate situation, but we need to distinguish this from non-anaphoric situational uses like the following:

- (81) Bring **that** along with you.

One only appears in singular count use, where it is near-obligatory in identity of sense use, optional in strict anaphoric use, and absent in non-anaphoric use. Now,

¹⁶ Hawkins's claim is that noun phrases like *we soldiers* are appositive structures, consisting of a pronoun with a full noun phrase in apposition with it. This is also the position of Delorme and Dougherty (1972), who see *we men* as identical in structure with *we, the men*, the only difference being that the appositive noun phrase is indefinite, with the zero determiner normal in plural indefinites. They point out that Postal's account fails to explain the impossibility of singular **I boy*, which falls out naturally on their analysis since *boy* is not a possible singular indefinite noun phrase. But their own analysis fails equally to account for the impossibility (in standard English) of *they men*.

the forms without *one* – *this*, *that*, *these*, *those* – in non-anaphoric and strict anaphoric use can only have non-human reference (while *these/those* can be human in identity of sense use). Human reference in sentences like (79) and (81) would be made by stressed *him*, *her* or *them*. The claim is that these pronouns used in this way are forms of a demonstrative. Further evidence comes from the following paradigms; the noun phrases are to be understood as non-anaphoric, and the capitalized forms as stressed:

- (82) a. Take a look at the car/teacher.
 b. Take a look at him.
 c. Take a look at her.
 d. Take a look at it.
 e. Take a look at them.
- (83) a. *Take a look at the car/teacher.
 b. Take a look at him.
 c. Take a look at her.
 d. *Take a look at it.
 e. Take a look at them.

THE and *IT* here are replaced by *that* (or *this*). Moreover, in standard English, *THEM* must be human, and for non-human reference would be replaced by *those* (or *these*). There is thus a neat complementary distribution between third-person personal pronouns and demonstratives, the latter filling certain gaps in stressed occurrences of the former.¹⁷

The conclusion is that the third-person pronouns *he*, *she*, *it*, *they*, when unstressed, represent pronominal forms of the definite article, but in addition some of them, when stressed, also represent forms of a demonstrative. Since they do not encode any degree of distance or proximity to any person, these demonstrative forms can be regarded as belonging to the deictically unmarked, or perhaps [– Prox], demonstrative already identified for English in 1.2.2. The forms of this demonstrative are, then, in the third person: singular pre-nominal *that*; singular pronominal *he*, *she*, *that*; plural pre-nominal *those*; plural pronominal *they* (human), *those* (non-human). Turning to first- and second-person pronouns, it is natural to conclude that the stressed forms of these are also demonstrative – first- and second-

¹⁷ *The* and *it* are both capable of bearing stress, to express contrast, for example, as in *I'm THE owner of this house, not one of the owners* and *I just met Julia walking her poodle; SHE looked fed up, but IT was chirpy enough*. What is impossible is the kind of demonstrative stress likely to be accompanied by pointing or some equivalent gesture appropriate to (83).

Among other qualifications to the pattern observed is the fact that *this* and *that* do occur with human reference in subject position in copular sentences: *Who's this?*, *That's Jill's husband*. But this exception may be only apparent because *it* is used in the same way, suggesting that a demonstrative used in this way is neuter: *Who is it?*, *It's Jill's husband*.

person forms of the non-deictic demonstrative – while the unstressed forms are simple definite. And the demonstrative-like behaviour of pre-nominal *we* and *you*, shown in (78), is straightforwardly accounted for; these too are demonstrative, the corresponding third-person form being *those*.¹⁸ I take personal pronouns generally, then, to be determiners. When unstressed they are forms of a definite article, and in some, but probably not all, stressed occurrences they are forms of a demonstrative.

What distinguishes personal pronouns from full definite (or demonstrative) noun phrases is their lack of descriptive content (beyond partly descriptive grammatical features like gender in some forms). A pronoun is therefore used, in general, where the associated descriptive content can be readily recovered from the discourse or the non-linguistic context. But there is much more to pronouns than this, and their grammar is enormously complex, involving some of the most central issues in syntactic theory. The theory of **binding** sets out the anaphoric possibilities of different types of noun phrase, dividing these into three categories accounted for by three binding principles. Reflexive pronouns (like *myself*, *themselves*) and reciprocals (like *each other*) are handled by Principle A, which says that these expressions (labelled **anaphors**) must be “bound” by an antecedent, and that the anaphor and its antecedent must both appear within a particular syntactic domain, called the anaphor’s “governing category”. The governing category is typically the minimal clause, but it may sometimes be a noun phrase or other phrase, and can be a complex clause; in the simplest cases, an anaphor and its antecedent must be in the same minimal clause (thus *The gangster shot himself*, but **The gangster is afraid the police will shoot himself*). Ordinary personal pronouns are the second category of nominal expression, and Principle B says they may be bound by an antecedent but need not be, and if they are the antecedent cannot be in their “governing category”. Thus, in *The gangster shot him*, the pronoun *him* may take its reference from some noun phrase occurring earlier in the discourse, but *the gangster*, which is in the same minimal clause, cannot function as its antecedent. Principle C is concerned with all remaining referential noun phrases (full noun phrases essentially), and these are not subject to binding. We have seen, of course, that definite full noun phrases may be anaphoric, but it is important to realize that binding is a technical concept defined partly in terms of structural configuration;

¹⁸ There are gaps in the paradigms of the deictically neutral or [– Prox] demonstrative and of the definite article, for these plural forms as well as for *I* and singular *you*. *We* cannot occur in the exclamatory use: *You fools!* but **We fools!*. And in non-exclamatory uses the pre-nominal simple definite article only occurs in the third person: *The linguist(s)* but (with unstressed pronunciation of the determiner) **I/you/we linguist(s)*. By contrast, in their pronominal use *I*, *you* and *we* show the full range of occurrence available to the third-person forms, and can appear with full or reduced pronunciation fairly freely.

it does not just mean “having an antecedent”.¹⁹ For an accessible account of the binding theory, see Haegeman (1994: chapter 4).

Work on the semantics of pronouns has identified further complexities in their behaviour which might be taken to argue against their being merely determiners without the descriptive content found in full noun phrases. As well as the straightforward referential uses of pronouns in which they get their reference either from a linguistic antecedent or from the situation, we can distinguish the **bound-variable** use, the **E-type** use, and the **lazy** use. For discussion of these and of the semantics of pronouns more generally see Cooper (1979), Hausser (1979), Evans (1980), Reinhart (1983). Bound variable pronouns are dependent on (or “bound” by) a quantifying expression (such as one expressible in terms of the logician’s universal quantification), and, though singular, do not have a specific referent but rather denote a range of individuals. For example:

(84) **Every girl** thinks **she** should learn to drive.

(85) **Every student** thinks **they** have passed the exam.

Note that *they* in (85) is the vague singular use, becoming increasingly common nowadays where the antecedent is of mixed gender. These examples are anaphoric in that the pronoun has an antecedent, *every girl* and *every student*, but this antecedent defines a range of entities and the pronoun refers to each of these individually. E-type pronouns are similar, but the antecedent of the pronoun is itself a variable bound by a quantifier, and the configurational relationship between pronoun and antecedent is not the one defining binding (because the antecedent does not **c-command** the pronoun: for explanations of c-command see Radford (1988: chapter 3) and Haegeman (1994: chapter 2)):

(86) **Every man** who bought **a car** crashed **it**.

Lazy pronouns are so called because they relate somewhat sloppily to an antecedent, agreeing with it in descriptive content rather than referential identity:

¹⁹ It is nevertheless important not to overstate the difference between personal pronouns and full definite noun phrases. There is in fact a class of full definite noun phrases sometimes termed “epithets”, typically with affective content showing something of the speaker’s attitude to the referent, with behaviour resembling that of third-person personal pronouns:

- (i) Finally **John** arrived, and you could see **he** was pleased with himself.
- (ii) Finally **John** arrived, and you could see **the lad/bastard/little darling** was pleased with himself.

The inherent definiteness of personal pronouns may, in fact, be an important factor in their behaviour as regards binding.

- (87) I keep **my car** in the garage but my next-door neighbour keeps **it** in his drive.

The pronoun *it* here refers, of course, to my neighbour's car, not mine. But most of these uses are not, in fact, restricted to pronouns. Kempson (1988) shows that full definite noun phrases can have the bound-variable and E-type uses:

- (88) Of **every house** in the area that was inspected, it was subsequently reported that **the house** was suffering from subsidence problems.
- (89) **Everyone** who bought **a house** discovered too late that **the house** was riddled with damp.

The lazy use seems to be a purely pronominal phenomenon,²⁰ but otherwise personal pronouns are strongly parallel in behaviour and range of use to definite full noun phrases. The claim stands that pronouns are definite noun phrases minus the description – thus determiners.

1.2.6 *Universal quantifiers*

Determiners like *all*, *every* and *each* can be thought of as approximating to universal quantification in logic. In fact these determiners differ in important ways from the logician's universal quantifier, and from each other, but this need not concern us here. For discussion see McCawley (1981: 98–101). It seems obvious that the behaviour of these expressions in the diagnostics in 1.2.1, aligning them with definites, is to be related to inclusiveness, since they express totality, either within a context or absolutely.

It is to be noticed, however, that some other determiners which do not express totality show similar behaviour to *all* etc. in at least some of the diagnostic environments:

- (90) a. Strong as **most** contestants are, they can't shift it.
b. **Most** hats are yours.
c. ?Are there **most** visitors here?

²⁰ Interestingly, in Kempson's discussion this use, peculiar to pronouns, is paralleled by a use peculiar to full definite noun phrases, the bridging cross-reference or associative anaphoric use discussed above in 1.1.2. Kempson does not suggest this, but it is tempting to argue that lazy pronouns are the pronominal equivalent of bridging cross-reference. Examples like the following suggest a case can be made for this view:

- (i) I glanced into **the kitchen** and saw that **the windows** were filthy; in **the bathroom**, on the other hand, **they** were quite clean.

The pronoun is interpreted, in terms of description, on the basis of the preceding noun phrase *the windows*. But while *they* refers to windows, it does not refer to the same windows; this is what makes it a lazy pronoun. It gets its reference from association with *the bathroom*, just as *the windows* gets its reference from association with *the kitchen*. The suggestion is, then, that *it* in (87) is similarly an associate of *my next-door neighbour*. But, being a pronoun, it also needs a full noun phrase "antecedent" to permit recovery of the description to be assigned to its referent; *my car* serves this function.

What *most* has in common with *all* is that it expresses a proportion of some whole – as indeed does *the* if it is inclusive. This has led to the suggestion that the inclusiveness of the definite article is one case of a broader concept of **quantification**, characterizing determiners which denote a proportion. This is an issue I will take up in Chapters 6 and 7.

1.3 Simple and complex indefinites

The article *a* is the obvious signal that a noun phrase is indefinite, but just as definites need not involve the definite article, so there are indefinites which do not contain the indefinite article. We here consider the range of indefinite noun phrase types, and look more closely at the indefinite article itself and its role in the expression of indefiniteness.

1.3.1 Indefiniteness and cardinality

An important question is whether indefiniteness is a function of the presence of certain indefinite determiners in a noun phrase, or whether it is simply a matter of the absence of definite determiners. The indefinite article *a* suggests the former, on the assumption that *a* encodes [– Def]; and of course *a* and *the* are mutually exclusive. But the object noun phrases in the following sentences are also indefinite, involving neither identifiability nor inclusiveness:

(91) I bought **three books** this morning.

(92) I wonder if Helen has read **many books**.

These indefinite noun phrases do not contain the indefinite article however. So perhaps they should be considered to be complex indefinites, with indefiniteness being signalled by some other determiner. Indeed determiners like *three* and *many* which denote **cardinality** – that is, a number or an amount, as opposed to a proportion – have been characterized as indefinite determiners. But it is clear that they do not encode [– Def]. Unlike *a*, they can co-occur with definite determiners:

(93) Pass me **those three books**.

(94) I've only read a few of **the many books** she's written.

These cardinality terms are obviously neutral with respect to (in)definiteness. There are some determiners which only appear in indefinite noun phrases; *some* and *enough* are examples, and these could reasonably be categorized, together with *a*, as indefinite determiners. But what makes the noun phrases in (91)–(92) indefinite is the absence of any definite determiner. Notice also that count nouns in the plural and mass nouns can occur without any determiner, and are then (at least where not interpreted generically) indefinite:

(95) John has gone out to buy **milk**.

(96) I've already put **spoons** on the table.

The indefinite article *a* only occurs, in fact, in singular count noun phrases – an odd limitation – and this has led many grammarians to suppose that its place is taken in examples like (95)–(96) by a “zero” variant of the indefinite article. But notice that even singular indefinite count noun phrases do not necessarily take *a*: *one orange*. This is not because *one* is an indefinite determiner; it is a cardinality term like *three* and *many*, and, like them, can appear in definite noun phrases: *the one orange*, *this one ticket*. So the idea that there is in English a zero plural and mass indefinite article finds little support; the reality seems to be that a noun phrase is indefinite if it has no definite determiner, whether or not it has an indefinite determiner. Another determiner traditionally regarded as being a plural and mass variant of *a*, or at least as complementing *a*, is *some*. In fact it is necessary to distinguish stressed *some*, with a full vowel, from a form which is usually (though probably not always) unstressed and with reduced vowel, thus pronounced [səm] – and often represented for convenience as *sm* (a convention I shall adopt). *Some* can occur with singular count nouns (*There's some man at the door*), and it is *sm* which partially complements *a* and is reasonably taken to be an indefinite article. But again, plural and mass noun phrases do not need to take *sm* to be indefinite, as seen from (95)–(96) (and (91)–(92)).

In view of these facts, one must ask whether *a* and *sm* really are indefinite articles. They are not in full complementary distribution with definite determiners, since we can have noun phrases with neither; in fact *a* does not co-occur with either the definite article or with any numeral, though *the* and numerals can co-occur. There is in fact good reason to classify *a* and *sm* as cardinality expressions. *A* is derived historically from the same ancestral form as the numeral *one*, and while few writers now recognize a synchronic link, an exception is Perlmutter (1970), who argues that *a* is derived from *one* by a phonological rule of reduction operating in the absence of stress. I prefer to regard *a* and *one* as distinct items, but if they are assigned to the same lexical class (both being cardinality terms, perhaps even numerals), it is clear that there is a particularly close relationship between them.²¹

This position is strengthened by the fact that in many languages the semantic equivalent of *a* is identical to the equivalent of *one*: German *ein*, French *un*, Turkish *bir*. In these languages, although descriptive grammars do tend to make a distinction between an indefinite article and a singular numeral, they are mainly distinguished

²¹ There are many contexts in which they are interchangeable, with little difference of meaning: *a/one fifth*, *a/one mile away*, *a/one week from now*, not *a/one bit*. *A* is also numeral-like in only occurring with count nouns, though this is not true of *sm*, which aligns with vague cardinality terms like *much/many* rather than numerals.

by stress; it may be, therefore, that in these languages the “article” is merely the unstressed variant of the singular numeral. But the differences are more substantial in English, where *a* can be contrastively stressed and is then realized as [eɪ], not as *one*. I take *a*, therefore, to be a cardinality term, perhaps even a kind of numeral, but a distinct lexical item from *one*. I suggest that the semantic difference between them is that *one* is in contrast with *two*, *three* etc., while *a* is in contrast with ‘more than one’. *A* thus encodes no more than [+ Sg], while *one* has some additional content besides. *Sm* can be thought of as expressing a vague quantity, but (if the noun is count) more than one; it may therefore stand in opposition to *a*, encoding [– Sg] (this specification applying both to plurals and, redundantly, mass noun phrases). But whereas *a* obligatorily accompanies singular indefinites in the absence of some other determiner, this is apparently not so of *sm* and non-singular indefinites.²²

The treatment of *a* as a cardinality expression, closely related to *one*, while attractive, does leave unexplained the fact that it cannot co-occur with *the* or other definite determiners. Since numerals can follow these determiners, **the a house* or **this a car* should be possible. On Perlmutter’s view that *a* is simply the unstressed form of *one*, one might claim that this numeral can only be unstressed and reduced when initial in the noun phrase, so that after another determiner the full form is appropriate; thus *the one house*, *this one car*. But *one* is purely optional here, whereas in the absence of a definite determiner a cardinality expression is obligatory in the singular, with *a* as the default form (the one that must occur if no other does). So for the Perlmutter analysis the problem is to explain why some cardinality expression is required if there is no definite determiner but optional otherwise. For my analysis the problem is to explain why a particular cardinality term, *a*, is obligatory

²² The distribution of *sm* in relation to plural and mass noun phrases such as those in (95)–(96) is a complex and controversial matter. It has been claimed, by Carlson (1977), that *sm* is the plural correlate of *a*, and the “bare plural” (as in (96)) is something radically different – not an alternative indefinite plural. We will return to a more detailed discussion of this in Chapter 4, and for the present it will suffice to note that in some contexts a plural or mass noun phrase may occur either in bare form or with *sm* with little difference in meaning, whereas in other contexts there may be a more substantial difference. Consider the following:

- (i) I’m looking for a record.
- (ii) I’m looking for *sm* records.
- (iii) I’m looking for records.

Sentence (i) is ambiguous between a reading on which the speaker is looking for a particular record and one on which she does not have a particular record in mind. Pairs of readings of this sort, commonly available with indefinites, are termed **specific** and **non-specific**, respectively. Turning to the plurals, the same ambiguity appears in (ii). But (iii) is different; here, with the bare plural, the specific reading, that the speaker has particular records in mind that she hopes to find, is either unavailable or at least difficult to get.

with singulars in the absence of any other determiner but impossible after a determiner. *The* and *a* are both phonologically weak forms, and this shared characteristic can be taken to reflect their both being the unmarked or basic members of their respective classes; *the* is the basic definite determiner and *a* is the basic cardinality determiner. These basic items act as default forms. A definite noun phrase must normally contain a definite determiner, and *the* is the one that occurs in the absence of some other with more semantic content. Indefinite noun phrases are characterized by the non-occurrence of a definite determiner, but both definites and indefinites should normally contain, in principle, an expression of cardinality. The fact that this can be dispensed with much more easily in plural and mass noun phrases may be related to the fact that plurals do have an expression of cardinality in the plural inflection on the noun, and cardinality is intrinsically less central, if indeed relevant at all, to mass expressions. With singulars, since the noun carries no number inflection, a cardinality determiner is in principle required, and *a* is the default form. But I suggest the issue is complicated by a phonological constraint that weak forms can only occur initially in the phrase.²³ The result is that in definites the initial definite determiner blocks the expected appearance of *a* in second, cardinality, position.

The basic, unmarked nature of *the* and *a*, with their minimal semantic content ([+ Def] and [+ Sg] respectively), reflected in their phonological weakness and default behaviour, I shall take to be what defines the term **article**. *The* is, then, the **definite article**, and I adopt the label **cardinal article** for *a* and *sm*. Notice though that the suggested phonological constraint blocking the appearance of *a* and *sm* after another determiner ensures that these items only appear in indefinite noun phrases. For this reason they can be said to signal indefiniteness indirectly. An article that does this, while not itself encoding [– Def], I shall describe by the further label **quasi-indefinite article**.

1.3.2 *Complex indefinites*

I use the term “complex indefinite”, informally, to denote noun phrases in which some determiner other than one of the quasi-indefinite articles seems to compel indefiniteness and render a definite determiner impossible. Cardinality terms other than *a* and *sm* tend to be compatible with definiteness as well as indefiniteness. But there are some cardinality expressions which show either a strong preference for indefiniteness (*several*) or incompatibility with definiteness (*enough*). *Several* is rarely used in company with a definite determiner (*the several trees*, *those several visitors*), but such sequences are not impossible and it

²³ The major exception to this is that “pre-determiners” such as *all* and *half* can occur before articles: *all the way*, *half an hour*.

may be simply that the meaning of *several* is such that it occurs more commonly in indefinites. But *enough* is incompatible with definiteness (**the enough sugar*, **those enough cups*).²⁴

There are other determiners which are not self-evidently cardinality expressions and which occur only in indefinite noun phrases. Most prominent in this group are *some* (as distinct from *sm*) and *any*. These determiners are traditionally labelled “indefinite”, and indeed must be serious candidates for consideration as encoding [– Def]. We shall see that determiners corresponding closely to these are identifiable cross-linguistically, many languages showing indefinite expressions forming a cline of specificity. We have observed that English *a* and *sm* can be specific or non-specific, and we will discuss *some* below in relation to this distinction. But it is clear that *any* expresses a kind of extreme non-specificity, in which the speaker does not merely have no particular entity in mind corresponding to the description, but does not care what is taken to satisfy the predicate. To express specificity unambiguously English has *a certain*, but there seems to be no reason to see this as a separate determiner. Rather, the adjective *certain* compels the specific interpretation available anyway to *a*.

Determiners in English can be used pronominally as well as pre-nominally, and we have seen in 1.2.5 that there can be substantial difference in shape between pronominal and pre-nominal forms. The pronoun corresponding to *a* is *one*, as illustrated by the following:²⁵

²⁴ It is not obvious why this is so. A synonym such as *sufficient* does seem to be acceptable in a definite context (though barely), which suggests that the incompatibility is not entirely semantic. There may be some syntactic constraint peculiar to the item *enough* behind the restriction. This determiner is unique in being able to follow the noun (*enough money* or *money enough*), a characteristic shared by its German equivalent *genug*. Now *genug* (which is equally excluded from definite noun phrases) also has the peculiarity of being uninflected. There are some other cardinality determiners in German, such as *viel* ‘much’, ‘many’ and *wenig* ‘little’, ‘few’, which can occur pre-nominally either with or without the normal inflection: *viel-er Kummer* or *viel Kummer* ‘a lot of trouble’, *wenig-e Freude* or *wenig Freude* ‘little joy’. But these cardinality terms must be inflected if preceded by a definite determiner: *der viele Kummer*, **der viel Kummer*; *die wenige Freude*, **die wenig Freude*. So the impossibility of *genug* after a definite determiner may be linked to its inability to take inflection. And this suggests the incompatibility with definiteness of *enough* may be an idiosyncratic syntactic property of this lexical item, left over from a constraint (similar to that in German) applying in an earlier stage of English when adjectives were inflected. This is speculation, and I will not pursue the point.

²⁵ Note that the *one* in (97) is distinct from that in (i):

- (i) I’m looking for a white hydrangea but I can only find a pink **one**.

Here a *pink one* is a complete noun phrase, with *a* as determiner and *one* as the head noun. This is the same minimal noun as in *the blue one* or *the one who Lucille divorced*, discussed above, and clearly has nothing to do with *a*. But if the *one* in (97) were this same minimal noun we would expect it to be preceded, like any count noun, by the article *a*; and this is impossible.

- (97) I'm looking for a white hydrangea but I can't find **one**.

Turning to *some* and *any*, these determiners, together with *sm*, can function pre-nominally and pronominally with no change in form, except that the pronominal form of *sm* is unreduced, thus *some*:

- (98) a. Joan met **some** man today.
b. Could you pass me **sm** broccoli, please?
c. I don't want **any** noise.
- (99) a. Give me **some**.
b. You haven't taken **any**.

But there are restrictions. Pronominal *some* can only be mass or plural, like *sm*; thus the occurrence in (99a) cannot be interpreted as singular count; pronominal *any* also tends to be mass and plural. Pronominal *some* and *any* are restricted to identity of sense anaphoric use.²⁶ In other functions they are replaced by fuller forms, *someone/somebody* and *anyone/anybody* for human reference (singular only), *something* and *anything* for non-human reference.²⁷

But while *anyone* and *anything* clearly correspond as pronouns to the determiner *any*, it is not obvious whether *someone* and *something* correspond to *some* or *sm*. The fact that *someone* is singular count suggests it includes *some*, not *sm*; but there is a striking difference in interpretation between *someone* and *some*. *Some*, with a singular count noun, implies that the referent is unfamiliar to the speaker, thus non-specific, as the following make clear:

- (100) a. There's some book lying on the table.
b. ?I've spent the afternoon reading some book.

But *someone* and *something* show the same ambiguity with regard to specificity as *a* and *sm*:

²⁶ Sentence (97) illustrates identity of sense anaphora (see 1.2.5 above), which is a major use of indefinite pronouns. The anaphoric expression relates to an antecedent with which it shares descriptive content, but the two are not coreferential. Note that with identity of sense anaphora the "antecedent" can be provided by the non-linguistic context, as when Mary is silently offered cake and says *Yes, I'll have some* or *I don't want any, thanks*. This, of course, is not really anaphora, but rather a kind of immediate situation use; I will, however, continue to use the term loosely to cover such cases.

²⁷ These complex forms combine the determiner with a noun, but they cannot be regarded as full, non-pronominal, noun phrases which happen to be represented orthographically as single words. This is clear from the position of adjectives: *someone important*, not *some important one* (which would only occur as an identity of sense anaphoric use of *some*, not necessarily human); and *something big* as distinct from the equally possible but not synonymous full noun phrase *some big thing*.

- (101) a. I'm looking for someone/something – but she/it doesn't seem to be here.
 b. I'm looking for someone/something – but I can't find anyone/anything I like.

Yet while *sm* is mass or plural, *someone* is singular, and *something* is singular or mass. So these pronouns do not correspond closely to either *some* or *sm*. They seem to correspond in fact to *sm* and *a* jointly. These articles between them cover the range singular, plural and mass; and they show the ambiguity between specific and non-specific. The fact that *someone* and *something* cannot be plural can be attributed to the incorporated noun.

The main uses of *any* and the pronouns based on it are illustrated by the following:

- (102) a. You can borrow **any** book you wish.
 b. **Anyone** could paint a picture like that.
 c. I'm not choosy; I'll eat **anything**.
 (103) a. Have you bought **any** vegetables?
 b. I don't want **any** wine; I've drunk enough.
 c. Have you had **anything** to eat?
 d. I don't think John's in his room; at least, I can't hear **anyone**.

Sentences (102) illustrate what I shall term the “random” sense, indicating an unrestricted choice. Examples (103) show the use of *any* in interrogative and negative contexts (which we can term jointly “non-assertive” contexts), where it tends to be preferred to *sm*. Non-assertive *any* tends to be unstressed and may be reduced (with pronunciations like [nɪ] or [nɪ̃], and similarly [nɪθɪŋ] etc.), while random *any* has to have some degree of stress and is phonologically full. But *sm* and related forms are also possible in non-assertive contexts, and, in particular, are perfectly good in interrogative sentences like (103a) and (103c). This suggests that non-assertive *any* may represent no more than a less emphatic use of random *any* rather than a distinct sense. On the other hand the two uses distinguished behave differently with regard to the diagnostic environments for definiteness; I illustrate this from the *there is/are* construction:

- (104) ??There is/are any book(s) in this library.
 (105) a. Are there any books in this library?
 b. There aren't any books in this library.

Any is perfectly good in non-assertive sentences with *there is/are*, but in positive declarative sentences, like (104), where it must be given the random interpretation, it is of very low acceptability. This seems to indicate that only non-assertive *any* is indefinite, and random *any* definite, which would presumably mean we have two

distinct senses, if not two distinct homophonous items. And this distinction correlates with another observation. *Any* seems to differ in position from cardinals in being able to precede numerals (*any three books, any one man*); but in this position (which could be definite determiner position) it can only have the random interpretation. So it may well be that random *any* is a definite determiner, occurring in the same “slot” as definites like *the* and *that*, while non-assertive *any* is a cardinality term, a variant of *a* and *sm* restricted to negative and interrogative contexts, occupying the same position in the noun phrase as other cardinals.

The discussion so far seems to leave the hypothesis that no determiner is specified [– Def], and that indefiniteness amounts to absence of a [+ Def] determiner, still plausible. Let us examine one further determiner which is a candidate for indefinite status, in that it never modifies definite noun phrases. *Such* has two principal uses, illustrated in the following sentences:

- (106) a. Mary is such a clever girl.
- b. I was given such lovely flowers.
- (107) a. We don’t need such a man here.
- b. I’d love to have such colleagues.

Whatever the syntactic status of *such* in (106), it is (on one interpretation) semantically a variant of *so*, and modifies the adjective. But in (107) it modifies the noun (an interpretation also possible in (106)), meaning something like ‘of that kind’ or ‘like that’. We need consider only the second use, which, it will be noticed, involves a demonstrative element.²⁸ We could simply say that *such* means ‘of that kind’ or ‘of this kind’, so that demonstrativeness is one element in its meaning. But it is interesting to observe at this point that noun phrases generally can have a “variety interpretation”. Consider the following sentences:

- (108) a. I wish I could afford to buy that car.
- b. I don’t like those boots.

As well as comments about a particular car and a particular pair of boots, these sentences could be used to refer to a type of car and a style of boots exemplified by the entity indicated (thus equivalent to *that type of car* or *a car of that type*, and *that style of boot* or *boots of that kind*). Now notice that *a car of that type* and *boots of that kind* are indefinite, and, moreover, very close in meaning to *such*

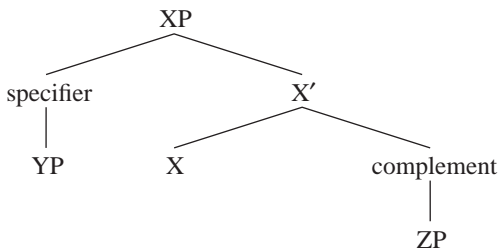
²⁸ *Such* is almost certainly an adjective, and a complete adjective phrase. It follows most cardinality expressions (*one such person, three such people*), but behaves like adjectives accompanied by degree modifiers in preceding the article *a*: *such a man; so/how/too clever a man*. The fact that *a* appears here in second position, which we have seen it is generally unable to do, suggests strongly that *such* and *so clever* etc. come to be in pre-article position as a result of a movement process, having originated in normal post-article adjective position.

a car and *such boots*. So *such* is synonymous, or near-synonymous, with a variety interpretation of a demonstrative. But perhaps one can go yet further, and propose that *such* is an indefinite demonstrative. On this proposal, *such* is synonymous with *this/that* (though not showing the proximal–distal distinction), but *this* and *that* occur only in definite noun phrases while *such* is the demonstrative form used in indefinite noun phrases. This would mean that demonstrativeness is not, as I have so far assumed, inherently definite. Interesting as this suggestion is, we will see in Chapter 3 that it does not seem to be supported by cross-linguistic evidence. It may, however, merit further investigation.

1.4 The noun phrase

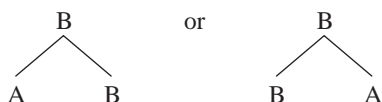
No investigation into the nature of definiteness can proceed far without consideration of the place of articles and other determiners within noun phrase structure. The question of how noun phrases are structured will be examined in more depth in Chapter 8, but it will be useful to establish some preliminary notions at this point.

The general view within generative syntactic theory is that the structure of phrases is determined by the principles of X-bar theory, according to which a phrasal category, an XP, is projected from a **head** X. Between these two is an intermediate level X'. XP immediately dominates, besides X', the position of **specifier** (which is therefore sister to X'). And the head may take one or more **complements** as sister, depending on its lexical properties. The specifier and complement positions are occupied by phrasal categories. This gives the following schema, in the linear order generally appropriate for English; linear order varies from language to language, and superficial (S-structure) order may differ from the underlying order of D(EEP)-structure.



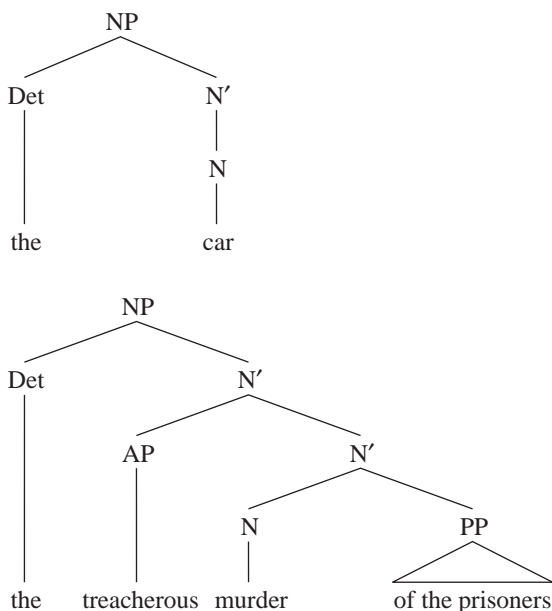
In addition, it is possible for other expressions, typically phrasal categories, to be **adjoined** to some or all of the projections of X (X itself, X', and XP). If A is adjoined to B, it is attached to it in this way:

Basic observations



For example, modifying expressions such as adjective or adverb phrases have been taken to be adjoined to X' or XP in the phrases they modify.

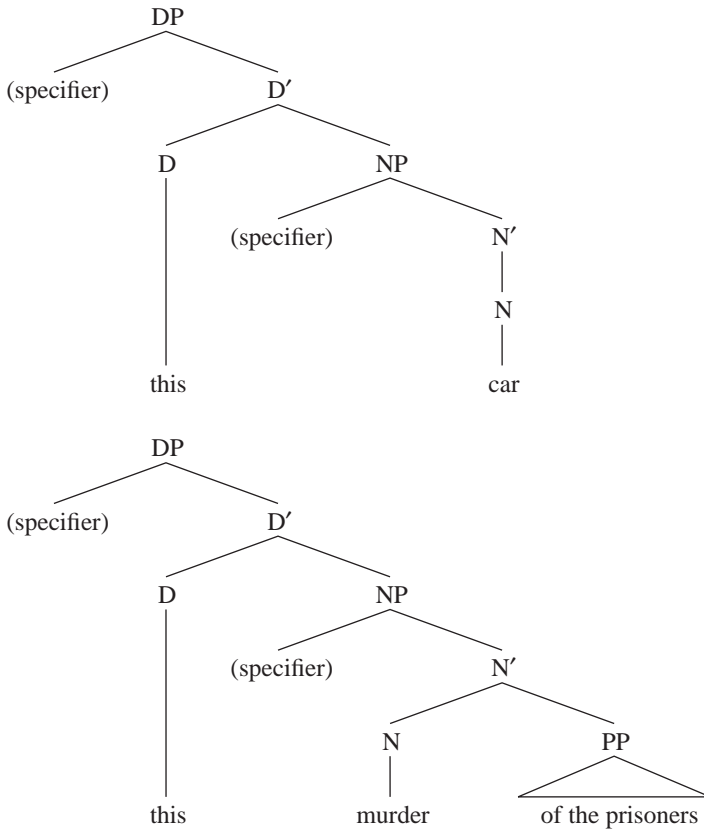
Given this general framework, the usual view until recently was that the noun phrase is a maximal (that is, phrasal) projection of the head N, with determiners in specifier position, adjectival expressions adjoined mainly to N', and any complements typically expressed by prepositional phrases. Some examples will make this clear:



For a clear and much fuller exposition of the general principles of X-bar theory and NP structure as just outlined, the reader is referred to Radford (1988: chapters 4 and 5).

A different account has recently come to prominence, and achieved almost general acceptance. This is that the "noun phrase" is a phrasal projection, not of the noun, but of the determiner (Det or D). It is therefore DP, not NP. There still exists a category NP, projected from N, but this is within DP, as complement of the head D, and corresponds more or less to N' on the older analysis. The general principles

of X-bar theory still apply, though it is now commonly assumed that adjunction to X' is not permitted. The following exemplify DP structure:



The position of adjectives has been the subject of much debate. They have been claimed to be adjoined to NP, but a currently influential view is that they are specifiers of some phrasal category between NP and DP.

The DP analysis arises from an interest in “functional” categories. These are grammatical categories as opposed to “lexical” or “substantive” categories like N(oun), V(erb), A(djective), P(reposition) (which can be thought of as denoting real-world entities, states, activities, properties, relations etc.). In clause structure, functional categories like T(ense) and Agr(eement) (representing subject–verb agreement) are treated as heads of phrases, and the DP analysis extends this trend to the “noun phrase”, taken to be a projection of the functional category of D(eterminer). Other functional heads, each projecting a phrase, have also been claimed to be involved in nominal structure; examples are Num(ber) and K (for case).

The question of which is the correct framework for nominal expressions will be taken up in Chapter 8 (where I will propose a substantially modified version of the DP hypothesis). For the present I will use the label *Det* to denote the formal category to which the English definite article and other determiners belong, but will leave open the question of whether it is the head or the specifier of the “noun phrase”.²⁹ The position of the definite article in English I will term non-committally “*Det position*” – which is either specifier of NP or head of DP depending on the framework chosen. The overall phrase, which is formally either NP or DP depending on the framework, I shall continue to term informally “noun phrase”.

English demonstratives and possessives apparently occupy the same slot in the noun phrase as *the*, and the definite quantifiers *all*, *both*, *every*, *most* tend to occur in this position too. *All* and *both* are sometimes termed “pre-determiners” because they may appear before definite Dets, to express a proportion of the whole denoted by the definite noun phrase: *all the girls*, *both your friends*. The phrase is equivalent to a partitive, and indeed an overtly partitive construction is available as an alternative: *all of the girls*, *both of your friends*. For reasons which are far from clear, *all* and *both* (at least in the pre-determiner construction) do not quantify over indefinites: **all sm girls*, **all a potato*, **both sm boys*. But they may occur with no following determiner:

(109) All three defendants were acquitted.

All here is probably in *Det*, rather than pre-determiner, position. *All three defendants* does not denote the totality of an indefinite group of three defendants, but of a definite group; it is synonymous with *all the three defendants*, not *all of three defendants* (if this is possible). Given the absence of *the*, the definiteness of the domain of quantification must come from *all*. The simplest explanation is that *all* is in *Det position*; when in this position it allows the noun phrase to be definite without *the*. *Most* has been argued to be a definite *Det*, and it is probably in *Det position* in *most men*, though it could be claimed to be in a cardinality position deeper inside the noun phrase, since it co-occurs neither with definite Dets nor with cardinality expressions. *Most* can also head a partitive construction, but cannot be a pre-determiner: *most of those men*, **most those men*. *Every* clearly occupies the same position as *the*; it cannot occur with it either as pre-determiner or (normally) in some more interior position, and it can precede numerals: **every the chair*, **the every chair*, *every one chair*, *every three chairs*. The only complication here is that it can, somewhat marginally, follow possessives: *his every whim*, *the king's every whim*. This is also possible after *the* but only with a possessive

²⁹ The convention has grown up of using the label *D* within the DP framework, while *Det* is preferred within the older NP framework. But the two are equivalent.

PP: *the every whim of the king*. So *every* seems to be capable of occurrence not only in Det position (the normal case) but also in some position deeper in the noun phrase, but only in the presence of a possessive expression; the marginality of this construction lies in its being available only with a limited range of head nouns: **his every car*, **the king's every castle*, **the every carriage of the king*. So it is possible to say that, in English, Det position is the locus not only for the definite article but for definite determiners more generally.

Most writers assume that *a* occupies the same position in noun phrases as *the*, but this is because they take *a* to be an indefinite article. If it is in reality a cardinal article, closely related to *one*, it is natural to suppose that it occupies (or at least originates in) the same position as numerals. Numerals agree closely in distribution with various other cardinality expressions – all being mutually exclusive and apparently occupying the same place relative to other constituents such as definite determiners (where they are compatible with these) and adjectives: *the one good book*; *the three/many/few good books*; *enough good books*. I shall assume that the English noun phrase has a position for cardinality expressions between the definite Det position and the position of adjectival modifiers, and that all the above words appear in this position.³⁰ A general similarity in behaviour and distribution between numerals and other cardinality determiners is to be found in many languages, as well as a strong tendency for these expressions to stand closer to the noun than definite determiners. I will assume, therefore, that syntactic theory makes available to languages a cardinality position which is more internal to the noun phrase than Det position.

1.5 Definiteness beyond the noun phrase

I end this chapter with the suggestion (which I will not pursue in the rest of the study) that definiteness is not only a feature of noun phrases, but occurs more widely. Let us consider two possible instances of it.

The tense–aspect distinction between past historic or preterite *I read that book* and perfect *I have read that book* has sometimes been described in traditional grammars in terms of a distinction between “definite” and “indefinite” past. These may be mere labels, but it is arguable that the preterite does make a definite time reference. In the absence of a time adverbial which identifies the time of the

³⁰ It may also be that they belong to the same category – possibly a phrasal category QP (quantifier phrase); this label is taken from Bresnan's (1973) study of the comparative construction, though Bresnan does not include numerals in her treatment. She defines *much*, *many*, *little*, *few* as possible heads of QP, which also contains a Det position; possible QP Dets are *as*, *so*, *-er* (*-er much/many* underlying *more*). If numerals and *a* are QPs, they are probably heads of category Q, differing from *many* in not admitting a Det. In this they are like *enough* and *several*. On the idea of a cardinality position less peripheral than definite Det position see also Verkuyl (1981).

event (*I read that book yesterday*), the hearer is assumed to be able to locate the event temporally on the basis of contextual knowledge. The perfect also presents the event as past, but, while the speaker may know when she read the book, there is no implication that the hearer knows or can work out (or needs to) when the event occurred. Put this way, the distinction is closely parallel to that between *the car* and *a car*, with identification of time of event substituted for referent identification.³¹

The structural position of determiners in the noun phrase is paralleled in the adjective phrase (and in adverb and quantifier phrases) by “degree modifiers” (Deg): *as/so/that/too big*. These words can be treated as being also of category Det, and, like this, as being either specifiers of AP or heads of a functional phrase containing AP. Note in particular that *that* operates both as a definite Det in the noun phrase and as a Deg in the adjective phrase, and that *this* also occurs in both uses:

- (110) a. Tom is stupid but not that stupid.
b. The fish I almost caught was this big!

There is little reason to doubt that *this* and *that* have demonstrative meaning in this use; the degree they convey of the property expressed by the adjective is accessed anaphorically in (110a) and communicated by means of an ostensive gesture in (110b), exactly parallel to what happens with noun phrase demonstratives. *This* and *that* as Degr are colloquial, and their more formal counterpart is *so*. As their near-synonym, *so* must also be a demonstrative Deg, lacking only the deictic distinction. If this is correct, the obvious question is whether there is then also a simple definite Deg. I believe the Deg to consider for this characterization is *as*. It is phonologically weak, with a normally reduced vowel, and its use in examples like the following is close to that of *the*:

- (111) a. Joe is as bright.
b. Joe is as bright as Ann.
c. Joe is not as bright as you think.

In (111a) the degree of brightness referred to is accessed by the hearer from the context or the preceding discourse, and in (111b) and (111c) it is provided in a relative-like modifier.

³¹ Partee (1984) discusses parallels, not exactly between tense and definiteness, but between tense and (definite) pronouns. She observes that, just as pronouns can relate to a referent introduced in the previous discourse or to a referent understood on the basis of the context, so tense can relate to an antecedent time or to an understood time.